JBS Journal of Bioscience Studies

Volume 2 | Issue 2 | 2025

Original Article

Open Access

Machine Learning Based Detection and Classification of Child Malnutrition Using Anthropometric and Demographic Data

Faisal Hameed[1], Atif Masih[2], Ahsan Balal Shaker[3], Syed Waqas Najeeb[4]

[1,3,4]Department of Computer Science, University of Engineering and Technology, Lahore, Pakistan

[2]Department of Computer Science, Government Islamia Graduate College, Kasur, Pakistan

Corresponding Author: Faisal Hameed

Email: fysalhameed@gmail.com

ABSTRACT

Background: Malnutrition among children remains a critical public health concern, particularly in regions with limited access to timely medical assessment.

Objective: This study presents a machine learning–based approach for the automated classification of child malnutrition using anthropometric and demographic data. The proposed system utilizes age, gender, height, and weight to predict nutritional status categories, including stunted, wasted, underweight, and overweight.

Methods: A publicly available malnutrition dataset was preprocessed through data cleaning, feature scaling, and class imbalance handling using oversampling techniques. Multiple supervised machine learning models, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and Gradient Boosting, were trained and evaluated. Model performance was assessed using accuracy, precision, recall, F1-score, and confusion matrices.

Results: Experimental results indicate that ensemble-based models, particularly Random Forest, achieve superior classification performance.

Implication: : This work contributes to the growing body of evidence supporting AI-assisted pediatric health screening and offers a reproducible baseline for future research. The findings demonstrate the potential of machine learning as a decision-support tool for early malnutrition screening; however, the system is intended solely as an aid and not a substitute for professional medical diagnosis.

Keywords: Malnutrition in children, Random Forest, Multiclass Classification, Data-driven healthcare

Received: 04-11-2025	Revised: 18-12-2025	Accepted: 24-12-2025	Published: 30-12-2025

How to cite:	Hameed F, Masih A, Shaker AB, Najeeb SW. Machine learning based detection and classification of child malnutrition using anthropometric and demographic data. J Biosci Stud. 2025;2(2):1-6. doi.org/10.65761/jbs.v2.i2.19
DOI	https://doi.org/10.65761/10.65761/jbs.v2.i2.19

INTRODUCTION

Malnutrition in children remains a major global pub-lic health challenge, particularly in low- and middle-income countries, where it contributes significantly to childhood morbidity and mortality [1]. According to global health reports, inadequate nutrition during early childhood is associated with impaired physical growth, weakened immune function with increased risk of illness and death, delayed cognitive development, and poor school performance in affected children [2]. Common forms of child malnutrition include stunting, wasting, underweight, and overweight, each of which reflects different nutritional deficiencies or imbalances and requires timely identification for effective intervention. A researcher states that conventional malnutrition assessment relies on anthropometric measurements such as age, height, and weight, interpreted using standardized growth charts and clinical expertise [3]. While these methods are clinically reliable, they are often time-consuming and dependent on trained healthcare professionals, which limits their accessibility in resource-constrained settings [4]. As a result, many malnutrition cases remain undetected or are identified only at advanced stages, reducing the effectiveness of treatment and intervention strategies. Recent advances in machine learning (ML) have demonstrated significant potential in healthcare applications, particularly in data-driven disease screening and classification tasks [5]. ML models are capable of identifying complex patterns in multidimensional data and can provide rapid, consistent predictions once trained [6]. In the context of child nutrition, anthropometric and demographic data represent structured numerical inputs that are well suited for supervised learning approaches. Several studies have explored ML techniques for health risk prediction; however, there remains a need for systematic evaluation of different classifiers for multi-class malnutrition categorization using readily available child health indicators [7, 8].

This study proposes a machine learning–based framework for the classification of child malnutrition using anthropometric data, including age, gender, height, and weight. Multiple supervised learning algorithms are trained and compared to identify the most effective model for predicting nutritional status categories such as stunted, wasted, underweight, and overweight. The approach emphasizes simplicity, interpretability, and practicality, making it suitable as a decision support tool for early malnutrition screening. It is important to note that the proposed system is intended to assist care-givers and healthcare practitioners and does not replace professional medical diagnosis.

The main contributions of this work are as follows:

•Development of a data-driven framework for multi-class child malnutrition classification using anthropometric features.

•Comparative evaluation of multiple machine learning models on a publicly available malnutrition dataset.

•Analysis of model performance using standard classification metrics to identify the most effective approach for early screening.

The remainder of this paper is organized as follows: Section II reviews related work on machine learning approaches for malnutrition detection. Section III describes the proposed methodology and data preprocessing steps. Section IV discusses the results and findings, while Section V concludes the paper.

RELATED WORK

Malnutrition assessment has traditionally been based on anthropometric indicators such as height-for-age, weight-for-height, and weight-for-age, interpreted using World Health Organization (WHO) growth standards [2]. These methods are widely accepted in clinical and public health practice; however, they require trained personnel and manual analysis, which can be challenging in resource limited environments. As a result, re-searchers have increasingly explored computational and data-driven approaches to support early malnutrition detection [9]. A recent analysis confirms the growing use of machine learning on demographic and health survey data for child malnutrition prediction across countries [10]. Comparable spatial machine learning models have been applied to childhood stunting in Pakistan [11]. Deep learning–based detection methods have also been explored, though their use remains comparatively limited [12], alongside machine learning–identified wasting risk factors in other national contexts such as Egypt [13].

[14] highlights the recent trends of Machine learning techniques that have been applied to various child health and nutrition problems due to their ability to analyze complex, multidimensional data. [15] reviews Several studies have utilized supervised learning algorithms to predict malnutrition outcomes using demographic and anthropometric features. Commonly used models include Logistic Regression, Decision Trees, Support Vector Machines (SVM), and ensemble methods such as Random Forests. These approaches have demonstrated promising results in identifying undernutrition patterns and classifying nutritional status with improved accuracy compared to traditional statistical methods. Similar frameworks applied to demographic and health survey data in Pakistan and Sub-Saharan Africa have achieved high predictive accuracy for stunting and wasting using Random Forest, SVM, and hybrid ensemble classifiers [16,17].

A researcher has focused specifically on binary classification problems, such as identifying whether a child is malnourished or not, using limited feature sets [18]. While these approaches provide useful screening capabilities, they do not distinguish between different forms of malnutrition, such as stunting and wasting, which require different intervention strategies. More recent research has addressed this limitation by adopting multi-class classification frameworks to predict multiple malnutrition categories simultaneously, thereby offering more informative clinical insights [19], including explainable machine learning approaches for identifying undernutrition predictors [20].

Class imbalance is a common challenge in malnutrition datasets, as certain nutritional categories may be underrepresented [21]. To address this issue, prior work has employed data resampling techniques such as Synthetic Minority Over-sampling Technique (SMOTE) [22] and cost-sensitive learning methods to improve model generalization and classification performance [23]. Studies have shown that balancing the dataset can significantly enhance recall and F1 scores for minority classes, which is crucial in health-related applications where false negatives can have serious consequences [24].

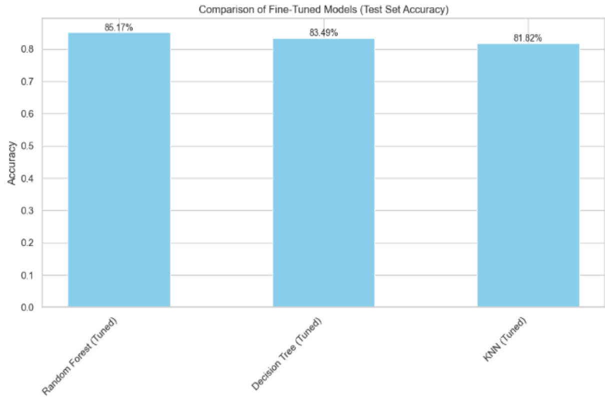
Despite these advancements, there remains a lack of comprehensive comparative analyses that evaluate multiple machine learning classifiers on the same anthro-pometric dataset using standardized evaluation metrics. Additionally, many existing studies emphasize model accuracy without sufficient discussion of interpretability and practical applicability. This work builds upon existing literature by systematically comparing several supervised machine learning models for multi-class malnutrition classification using easily obtainable anthropometric data, with a focus on early screening and decision-support rather than clinical replacement.

METHODOLOGY

This part will discuss the proposed machine learning architecture of child malnutrition classification, such as the description of datasets, preprocessing of data, selection of model, and the general classification process. The aim is to estimate the various types of malnutrition based on the easily available anthropometric and demographic characteristics.

Dataset Description

The dataset used in this study was obtained from a publicly available GitHub repository focused on malnutrition detection using machine learning. The dataset contains child-level anthropometric and demographic information, including age, gender, height, and weight. Each record is associated with a nutritional status label representing one of the malnutrition categories: stunted, wasted, underweight, or overweight. These categories reflect widely accepted child nutrition indicators based on standardized growth assessment practices. The figure 1 shows distribution of data against malnutrition categories.



Model	Accuracy	Precision	Recall	F1-Score
Random Forest (Tuned)	0.8516746411483254	85.54%	85.17%	84.83%
Decision Tree (Tuned)	0.8349282296650717	83.93%	83.49%	82.89%
KNN (Tuned)	0.8181818181818182	82.50%	81.82%	80.84%

Figure 1: Visualization of data against Malnutrition Category

To understand the dataset, we visualize the distribution of dataset by sex and income group. The figure 2 and 3 visualize the data.

Figure 2: Distribution of data by Sex and income group

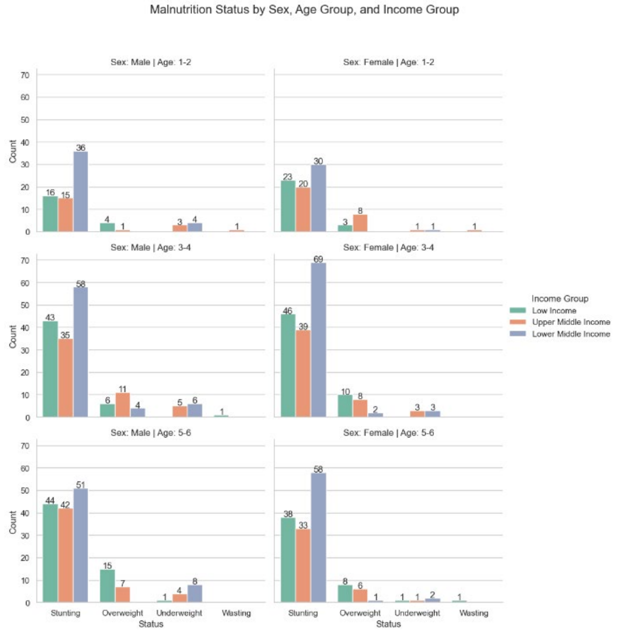


Figure 3: Distribution of child malnutrition status across sex, age groups, and income groups

The dataset includes samples from multiple nutritional classes, providing a suitable basis for supervised multi-class classification. However, as is common in real world health datasets, the distribution of samples across classes is imbalanced, necessitating appropriate preprocessing techniques prior to model training. Fig-ure 3 shows the individual feature distribution.

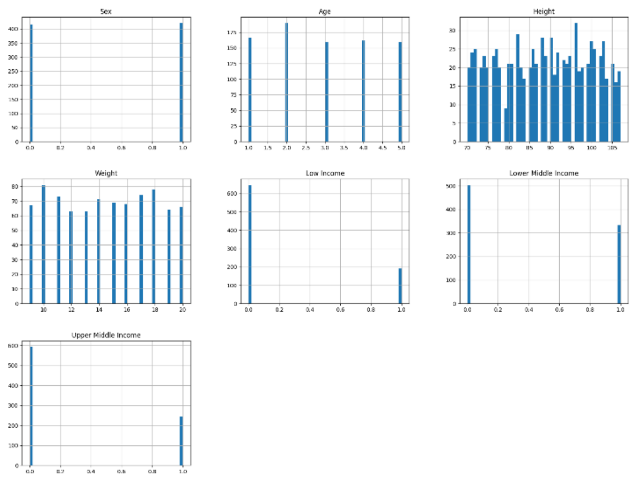


Figure 4: Distribution of child malnutrition status across sex, age groups, and income groups

Data Preprocessing

To ensure data quality and improve model performance, several preprocessing steps were applied. Missing values were examined and handled using appropriate imputation strategies to avoid information loss. Categorical variables, such as gender, were encoded into numerical form to make them compatible with machine learning algorithms. Continuous features including age, height, and weight were normalized using feature scaling techniques to prevent bias toward attributes with larger numeric ranges.

Outliers were identified and treated to reduce the impact of anomalous values that could negatively influence the learning process. To address class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to generate synthetic samples for underrepresented classes, thereby improving model generalization and reducing classification bias. Figure 5 shows class distribution before and after smote.

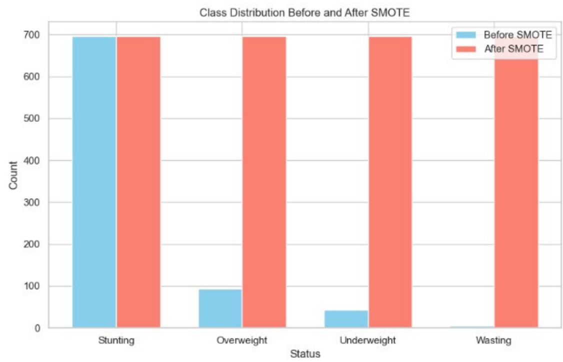


Figure 5: Class Distribution before and after SMOTE

To check the non-null values in each column, we visualize the data as in figure 6. Because no null values were there, so we just took this as it is.

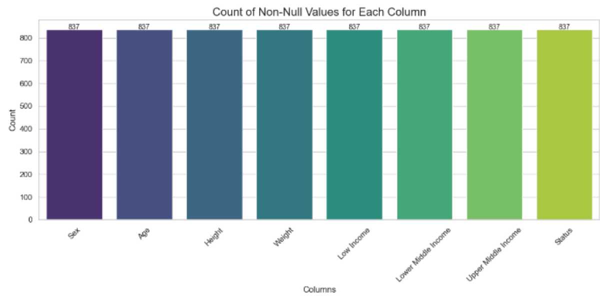


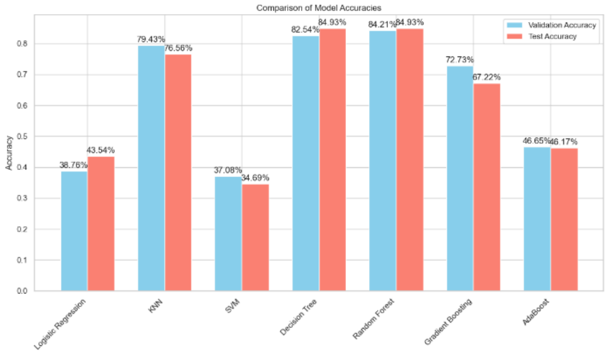
Figure 6: Class Distribution before and after SMOTE

Machine Learning Models

We employed multiple supervised machine learning algorithms to evaluate their effectiveness in classifying child malnutrition. These models were selected based on their robustness, interpretability, and proven success in healthcare related classification tasks:

- Logistic Regression (LR): This is a simple model that would be easy to interpret [25].
- Decision Tree (DT): This is a rule-of-thumb classification model and is capable of modeling non-linear relationships [26].
- Random Forest (RF): This is an ensemble learning algorithm which employs a set of decision trees to promote predictive accuracy and reduce overfit [27].
- Support Vector Machine (SVM): It can be applied to big-dimensional data, as well as in decision-making when the boundaries are complex [28].
- Gradient Boosting (GB): is an ensemble training that becomes more and more effective by focusing on the samples which have been misclassified.

Each model was trained on the preprocessed dataset and optimized using appropriate hyperparameters to achieve optimal classification results. The Comparison of results is shown in figure 7.



Model	Precision	Recall	F1-score
Logistic Regression	43.31%	43.54%	41.20%
KNN	79.93%	76.56%	74.25%
SVM	26.12%	34.69%	24.07%
Decision Tree	85.32%	84.93%	84.43%
Random Forest	85.31%	84.93%	84.56%
Gradient Boosting	67.16%	67.22%	65.95%
AdaBoost	46.07%	46.17%	45.02%

Figure 7: Model Comparison

Hyperparameter Tuning

Among the evaluated models, the Random Forest classifier demonstrated superior and consistent performance across all evaluation metrics. Consequently, it was further optimized using hyperparameter tuning techniques to improve generalization. Key parameters such as the number of trees, tree depth, and splitting criteria were adjusted to obtain the final optimized Random Forest model, which was used for subsequent analysis.

```
Tuned Random Forest - Test Set
Accuracy: 0.8516746411483254
              precision    recall  f1-score   support

  Overweight       0.77      0.85      0.81        93
    Stunting       0.88      0.68      0.76       120
 Underweight       0.83      0.94      0.88       113
     Wasting       0.94      0.98      0.96        92

    accuracy                           0.85       418
   macro avg       0.85      0.86      0.85       418
weighted avg       0.86      0.85      0.85       418
```

Figure 8: Random Forest Results

Ethical Considerations

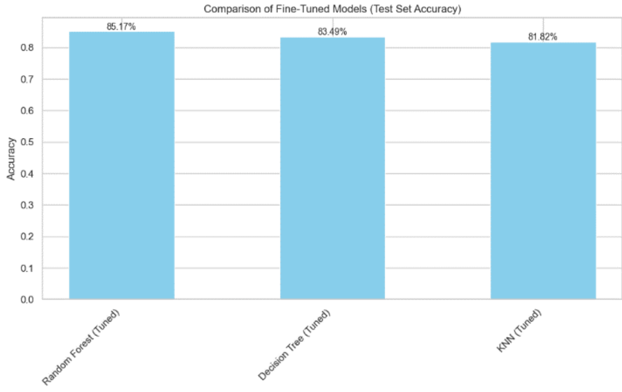
Even though personally identifiable information is not present in the dataset that was used in this study, ethical considerations are critical. The proposed machine learning system is aimed to facilitate the initial screening and education on the possible risks of malnutrition. It is not aimed at substituting professional medical examination or clinical diagnosis. The system needs to treat predictions as outputs of the decision support and verify them through competent healthcare.

RESULTS

In this section, the results of the experiments are given on the various machine learning models, and their performance compared with each other in classifying child malnutrition types using anthropometric data are discussed.

Performance Comparison of the Model

A comparative evaluation of the top three tuned machine learning models—Random Forest, Decision Tree, and K-Nearest Neighbors (KNN)—was conducted using accuracy, precision, recall, and F1-score, as summarized in Table X. Among all evaluated models, the tuned Random Forest classifier achieved the best overall performance, with an accuracy of 85.17%, precision of 0.855, recall of 0.852, and an F1-score of 0.848. These results indicate that the Random Forest model provided the most balanced trade-off between correctly identifying malnutrition cases and minimizing misclassification across all classes.



Model	Accuracy	Precision	Recall	F1-Score
Random Forest (Tuned)	0.8516746411483254	85.54%	85.17%	84.83%
Decision Tree (Tuned)	0.8349282296650717	83.93%	83.49%	82.89%
KNN (Tuned)	0.8181818181818182	82.50%	81.82%	80.84%

Figure 9: Comparison of Best three Models

The tuned Decision Tree model ranked second, achieving an accuracy of 83.49% and an F1 score of 0.829. While its performance was competitive, it exhibited slightly lower generalization capability compared to Random Forest, likely due to its sensitivity to data variations and limited ability to capture complex feature interactions. Nevertheless, its relatively high precision (0.839) and recall (0.835) demonstrate its effectiveness as a lightweight and interpretable classification model.

The tuned KNN model showed comparatively lower performance, with an accuracy of 81.82% and an F1 score of 0.808. Although KNN benefited from tuning, its reliance on distance-based similarity made it more susceptible to noise and feature scaling issues, which may have affected its performance on overlapping malnutrition classes. Figure 10 explains the comparison.

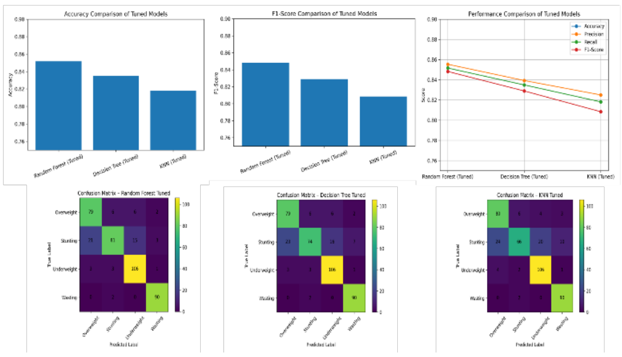


Figure 9: Comparison of best three Models

Overall, the results confirm that ensemble-based learning, particularly the tuned Random Forest model, is more effective for multiclass malnutrition classification using anthropometric data. Its superior performance across all evaluation metrics highlights its ro-bustness and suitability for deployment in early malnutrition screening systems. Consequently, the tuned Random Forest model was selected as the final classifier for this study.

DISCUSSION

The present study demonstrates that machine learning algorithms can effectively classify child malnutrition using readily available anthropometric and demographic variables.

Among the evaluated classifiers, the tuned Random Forest model achieved the highest overall performance, with an accuracy of 85.17%, precision of 0.855, recall of 0.852, and an F1-score of 0.848. These findings indicate that ensemble learning methods provide a robust balance between sensitivity and specificity, making them well suited for multiclass nutritional status classification. The superior performance of Random Forest is consistent with its ability to model complex nonlinear relationships while minimizing overfitting through the aggregation of multiple decision trees. This characteristic is particularly valuable in healthcare datasets, where interactions among age, sex, height, and weight are often heterogeneous and difficult to capture using conventional statistical approaches.

The results are consistent with previous studies reporting the effectiveness of machine learning in pediatric nutrition screening. Talukder and Ahammed reported that machine learning algorithms can accurately predict malnutrition among children under five using demographic and anthropometric characteristics. Similarly, Kar et al. demonstrated that supervised learning techniques substantially improve malnutrition prediction compared with traditional analytical methods. More recently, Saleem et al. highlighted the utility of Random Forest models for predicting child malnutrition in Pakistan, emphasizing their ability to identify complex risk patterns from demographic and socioeconomic data. The current findings further support these observations by demonstrating that ensemble-based models outperform individual classifiers in a multiclass classification framework.

An important methodological strength of this study is the application of SMOTE to address class imbalance before model development. Imbalanced datasets frequently reduce the ability of predictive models to recognize minority nutritional categories, resulting in lower recall and increased false-negative predictions. Balancing the dataset improved the discrimination of underrepresented classes and contributed to more consistent performance across evaluation metrics. Similar improvements following oversampling have been reported in previous studies investigating childhood nutritional classification, supporting the effectiveness of this preprocessing strategy.

From a practical perspective, the proposed framework offers several advantages for early screening programs. The required input variables, including age, sex, height, and weight, are routinely collected in community health services and primary healthcare settings, making the model inexpensive and readily deployable. Integration into digital health platforms or mobile decision-support systems could facilitate rapid identification of children requiring further nutritional assessment, particularly in underserved regions where access to pediatric specialists is limited. Nevertheless, the model should be considered a screening aid rather than a replacement for clinical assessment, as nutritional diagnosis requires comprehensive medical evaluation and consideration of socioeconomic, dietary, and clinical factors.

Despite these encouraging findings, several limitations should be acknowledged. The study relied on a publicly available dataset, which may not fully represent the diversity of pediatric populations across different geographical and socioeconomic settings. Furthermore, the prediction model incorporated only anthropometric and demographic variables, whereas additional clinical, dietary, environmental, and socioeconomic factors could further enhance predictive performance. Future studies should validate the proposed framework using larger multicenter datasets and explore explainable artificial intelligence techniques to improve model transparency and facilitate clinical acceptance. Such developments would strengthen confidence in artificial intelligence–assisted nutritional screening and support its integration into evidence-based pediatric healthcare.

CONCLUSION

In this paper, the authors have proposed a machine learning framework to classify malnutrition in children based on anthropometric measurements. The study used SMOTE to solve the issue of class imbalance and tested tuned classifiers across a variety of tunings, thus demonstrating the usefulness of datadriven methods in the process of early malnutrition screening. The tuned Random Forest classifier was the best performing model with accuracy equal to 85.17, precision equal to 0.855, recall equal to 0.852, F1-score equal to 0.848 indicating balanced and reliable performance across all the malnutrition categories.

The findings validate the claim that the ensemble-based models are highly appropriate in classification of multiclass malnutrition with simple anthropometric characteristics. Although the system is not meant to substitute clinical diagnosis but to aid it, it demonstrates a good potential in facilitating early detection of malnutrition especially in resource-depleted environments. The work in the future will be aimed at the expansion of data sets, expansion of features, and their application in practice.

Ethical Approval

Ethical approval was not required for this study because it was based exclusively on a publicly available, anonymized dataset containing no personally identifiable information. No human participants or animals were directly involved in this research.

Informed Consent

Informed consent was not required because the study analyzed publicly available, de-identified data and did not involve direct interaction with human participants.

Data Availability Statement

The dataset analyzed during this study is publicly available through a GitHub repository for child malnutrition prediction. Additional information regarding data preprocessing and analysis is available from the corresponding author upon reasonable request.

Author Contributions

Conceptualization: Faisal Hameed, Atif Masih; **Methodology**: Faisal Hameed, Atif Masih, Ahsan Balal Shaker; **Software**: Faisal Hameed; **Data Curation**: Faisal Hameed; **Formal Analysis**: Faisal Hameed, Ahsan Balal Shaker; **Writing – Original Draft Preparation**: Faisal Hameed; **Writing – Review & Editing**: Atif Masih, Ahsan Balal Shaker, Syed Waqas Najeeb; **Supervision**: Syed Waqas Najeeb. All authors have read and approved the final manuscript.

Funding

This research received no external funding.

Conflict of Interest

The authors declare that they have no competing interests.

Acknowledgments

The authors acknowledge the developers and contributors of the publicly available child malnutrition dataset used in this study.

REFERENCES

1. Dipasquale V, Cucinotta U, Romano C. Acute malnutrition in children: pathophysiology, clinical effects and treatment. Nutrients. 2020;12(8):2413.
2. World Health Organization. Malnutrition in children: consequences and implications [Internet]. Geneva: World Health Organization; [cited 2025 Jan]. Available from: https://www.who.int/data/nutrition/nlis/info/malnutrition-in-children
3. Becker PJ, Carney LN, Corkins MR, Monczka J, Smith E, Smith SE, et al. Consensus statement of the Academy of Nutrition and Dietetics/American Society for Parenteral and Enteral Nutrition: indicators recommended for the identification and documentation of pediatric malnutrition (undernutrition). J Acad Nutr Diet. 2014;114(12):1988-2000.
4. St-Jules DE, Lloyd L, Meade A, Biruete A, Kistler B, Carrero JJ. Deconstructing disease-related malnutrition: a new assessment framework for clinical practice. J Ren Nutr. 2023;33(6):707-16.
5. Talukder A, Ahammed B. Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. Nutrition. 2020;78:110861.
6. Jain S, Khanam T, Abedi AJ, Khan AA. Efficient machine learning for malnutrition prediction among under-five children in India. In: 2022 IEEE Delhi Section Conference (DELCON); 2022 Feb 11-13; New Delhi, India. Piscataway (NJ): IEEE; 2022. p. 1-10.
7. Kar S, Pratihar S, Nayak S, Bal S, HL G. Prediction of child malnutrition using machine learning. In: 2021 10th International Conference on Internet of Everything, Microwave Engineering, Communication and Networks (IEMECON); 2021 Dec 2-4; Bhubaneswar, India. Piscataway (NJ): IEEE; 2021. p. 1-4.
8. Constenla-Villoslada S, Liu Y, McBride L, Ouma C, Mutanda N, Barrett CB. High-frequency monitoring enables machine learning-based forecasting of acute child malnutrition for early warning. Proc Natl Acad Sci U S A. 2025;122(23):e2416161122.
9. Janssen SM, Bouzembrak Y, Tekinerdogan B. Artificial intelligence in malnutrition: a systematic literature review. Adv Nutr. 2024;15(9):100264.
10. Rao B, Rashid M, Hasan MG, Thunga G. Machine learning in predicting child malnutrition: a meta-analysis of demographic and health surveys data. Int J Environ Res Public Health. 2025;22(3):449.
11. Usman M, Kopczewska K. Spatial and machine learning approach to model childhood stunting in Pakistan: role of socio-economic and environmental factors. Int J Environ Res Public Health. 2022;19(17):10967.
12. Biradar VG, Naik KK. Detection of malnutrition in children using deep learning model. In: International Conference on Smart Computing and Communication; 2024 Jan. Singapore: Springer Nature Singapore; 2024. p. 35-45.
13. Hendy A, Abdelaliem SMF, Sultan HM, Alahmedi SH, Ibrahim RK, Abdelrazek EME, et al. Unlocking insights: using machine learning to identify wasting and risk factors in Egyptian children under 5. Nutrition. 2025;131:112631.
14. Jeganathan M, Jeganathan SR. Predictions of malnutrition among children. In: Transforming healthcare. Boca Raton: CRC Press; 2025. p. 77-98.
15. Indrisari E, Febiansyah H, Adiwinoto B. A systematic literature review on the application of machine learning for predicting stunting prevalence in Indonesia (2020–2024). Jurnal Sisfokom (Sistem Informasi dan Komputer). 2025;14(3):277-83.
16. Saleem MU, Aslam MU, Khatir AG, Jiang Q. Exploring risk factors and predictive modeling of child malnutrition in Pakistan using machine learning. Am J Hum Biol. 2025;37(9):e70124.
17. Khan MNA, Yunus RM. A hybrid ensemble approach to accelerate the classification accuracy for predicting malnutrition among under-five children in sub-Saharan African countries. Nutrition. 2023;108:111947.
18. Tamanna T, Mahmud S, Salma N, Hossain MM, Karim MR. Identifying determinants of malnutrition in under-five children in Bangladesh: insights from the BDHS-2022 cross-sectional study. Sci Rep. 2025;15(1):14336.
19. Worku MG, Mohanty I, Mengesha Z, Niyonsenga T. Magnitudes of various forms of undernutrition among children from the composite index of anthropometric failure in Sub-Saharan Africa: a systematic review and meta-analysis. Nutrients. 2025;17(11):1818.
20. Islam MM, Kibria NMSJ, Kumar S, Roy DC, Karim MR. Prediction of undernutrition and identification of its influencing predictors among under-five children in Bangladesh using explainable machine learning algorithms. PLoS One. 2024;19(12):e0315393.
21. Mgomezulu WR, Thangata P, Mkandawire B, Amoah N. Advancing predictive analytics in child malnutrition: machine, ensemble and deep learning models with balanced class distribution for early detection of stunting and wasting. Hum Nutr Metab. 2025;200340.
22. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321-57.
23. Najwa AF, Gunawan PH. Enhancing multiclass classification of child nutritional status using KNN and random forest with SMOTE. In: 2024 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA); 2024. Piscataway (NJ): IEEE; 2024. p. 483-8.
24. Gedefaw AE, Mengistu AK, Maru TC, Walle AD, Alemayehu MA, Yehuala TZ, et al. Predicting malnutrition in PLWHIV using machine learning in Gondar, Ethiopia. BMC Public Health. 2025;25(1):3177.
25. Hosmer DW Jr, Lemeshow S, Sturdivant RX. Applied logistic regression. 3rd ed. Hoboken (NJ): John Wiley & Sons; 2013.
26. Costa VG, Pedreira CE. Recent advances in decision trees: an updated survey. Artif Intell Rev. 2023;56(5):4765-800.
27. Salman HA, Kalakech A, Steiti A. Random forest algorithm overview. Babylonian J Mach Learn. 2024:69-79.
28. Pisner DA, Schnyer DM. Support vector machine. In: Machine learning. Cambridge (MA): Academic Press; 2020. p. 101-21.
29. Ayodele A. A comparative study of ensemble learning techniques for imbalanced classification problems. World J Adv Res Rev. 2023;19(1):1633-43.