



OPEN ACCESS

EDITED BY

Marta Esgalhado,
Lusófona University Research Centre for
Biosciences & Health Technologies, Portugal

REVIEWED BY

Leandro Oliveira,
Universidade Lusófona Research Center for
Biosciences & Health Technologies, Portugal
Luis Javier Sánchez Martínez,
Complutense University of Madrid, Spain

*CORRESPONDENCE

Priya Govindarajan
✉ priyagovindarajan@my.amrita.edu

RECEIVED 10 July 2025

ACCEPTED 22 September 2025

PUBLISHED 16 October 2025

CITATION

Shruthi S, Govindarajan P, Shalini SR,
Antony PJ, Uma AN and Rangarajan L (2025)
RISE: a novel unified framework for feature
relevance in malnutrition analytics integrating
statistical and expert insights.
Front. Public Health 13:1663373.
doi: 10.3389/fpubh.2025.1663373

COPYRIGHT

© 2025 Shruthi, Govindarajan, Shalini, Antony,
Uma and Rangarajan. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

RISE: a novel unified framework for feature relevance in malnutrition analytics integrating statistical and expert insights

S. Shruthi¹, Priya Govindarajan^{1*}, S. R. Shalini²,
Pavan John Antony³, A. N. Uma⁴ and Lalith Rangarajan⁵

¹Department of Computer Science, School of Computing, Amrita Vishwa Vidyapeetham, Mysuru, India, ²Pediatric Department, Mysore Medical College and Research Institute, Mysuru, India, ³School of Education, Adelphi University, Garden City, NY, United States, ⁴Genetic Unit, Anatomy Department, Mahatma Gandhi Medical College and RI, Sri Balaji Vidyapeeth (Deemed to be University), Puducherry, India, ⁵Department of Studies in Computer Science, University of Mysore, Mysuru, India

Addressing child malnutrition remains a critical global health priority, directly contributing to Sustainable Development Goals (SDG 2 – Zero Hunger and SDG 3 – Good Health and Well-being). This study aims to identify and prioritize the most influential determinants of acute forms of malnutrition among children aged 0–23 months by developing a novel feature scoring framework, RISE (Relevance-based Integration of Statistics and Expertise). The objective is to bridge the gap between data-driven modeling and context-specific insights by integrating model-based scores (from XGBoost), statistical filter methods for frequency boosting, and domain-informed adjustments. Using real-world data from Nutrition Rehabilitation Centre (NRC) at K.R. District Hospital, Mysuru, the RISE framework enhances the interpretability and contextual relevance of predictors often underweighted in traditional models. Domain-relevant features such as Mother Height, Breastfeeding Status, Caste, Maternal Working Status, and Ration card emerged as critical factors when adjusted through the RISE Framework. The top-ranked features included Child Weight, maternal anthropometry, and Child order remained consistently influential determinants, reflecting maternal dependency and the double burden of malnutrition. RISE uncovers hidden yet meaningful contributors that often go underrepresented in purely model-driven analyses. By adjusting feature scores to recognize both empirical strength and domain importance. By aligning analytical rigor with public health relevance, this study contributes a scalable, context-sensitive approach to feature prioritization in malnutrition research, supporting more informed, targeted interventions and policy actions toward achieving global nutrition goals.

KEYWORDS

child malnutrition, maternal malnutrition, feature scoring, model-based scoring, domain-based scoring, filter-based feature Scoring, frequency boosting, XGBoost

1 Introduction

Malnutrition remains one of the most persistent global health challenges, with every country striving to control, manage, and ultimately eradicate it. Malnutrition of any form, such as wasting, stunting, or underweight (1), can occur at various stages of life; its occurrence during childhood is considered the most alarming. Globally, stunting affected an estimated 23.2 percent or 150.2 million children, and wasting threatened the lives of an estimated 6.6 percent or 42.8 million children under 5 in 2024 (2). The outcome of the “Fifth National

Family Health Survey (NFHS-5)” discloses a prevalence of malnutrition in India, children under 5 years who are “stunted (height-for-age) 35.5, ‘wasted (weight-for-height)’ 19.3, ‘severely wasted (weight-for-height)’ 7.7, ‘underweight (weight-for-age)’ by the ‘World Health Organization’ (WHO) standards (3)”. Both national governments and international organizations have recognized its critical impact and continue introducing policies, programs, and guidelines to address this multifaceted issue. Addressing malnutrition during the formative years is a healthcare imperative and a socioeconomic priority. Malnutrition is influenced by region-specific, environmental, cultural, and socioeconomic factors (2–5). What contributes to malnutrition in one setting may not hold the same relevance in another, making context-sensitive analysis crucial (5). Policymakers prioritize the identification of malnutrition determinants before formulating or revising any policies or guidelines (4). Understanding these underlying factors is essential for effective intervention design, resource allocation, and impact measurement. As a result, the study of determinants of malnutrition has emerged as a critical area of research.

In recent years, machine learning (ML) has emerged as a powerful tool for identifying malnutrition determinants. Conventional feature selection methods, such as statistically based filter methods, have been widely used in ML to identify relevant features by evaluating statistical relationships between features and the outcome. These techniques are computationally efficient and easy to interpret, making them valuable tools in the early stages of exploratory data analysis (6–9). However, they often evaluate features independently, ignoring potential interactions among variables. These methods are sensitive to data distribution, scale, and noise, and may overlook contextually important features that do not exhibit strong statistical signals. This limitation becomes critical in complex public health problems like malnutrition, where domain knowledge and latent factors play a significant role. Hence, relying solely on conventional techniques may lead to the exclusion of key features that are crucial from a policy or intervention perspective (10–14). The research objective is to develop and implement the RISE (relevance-based integration of statistics and expertise) framework to enhance the identification of key determinants of acute forms of child malnutrition. By integrating model-based feature importance, domain knowledge, and frequency boost based on statistical relevance, the study aims to capture both dominant and overlooked factors that traditional statistical and machine learning models overlook. The RISE balances predictive power with contextual significance. Computationally, it introduces frequency boosting using a filter-based ensemble model, normalization across scales, and implementation of nested grid search for hyperparameter tuning. These positions RISE as both a technically rigorous and domain-sensitive framework, filling a critical gap in malnutrition-related ML research.

1.1 Literature review

The application of machine learning (ML) to malnutrition prediction has evolved significantly over the past few years. Early work by Anku et al. (7) demonstrated the superiority of XGBoost (98% accuracy, 100% AUC) in predicting wasting, stunting, and underweight in Ghana. Talukder et al. (8) demonstrated the potential of ML algorithms such as random forest (RF), logistic regression, and

k-nearest neighbors (k-NN) in identifying malnutrition among children, with RF achieving the highest sensitivity (94.66%) and specificity (69.76%). In 2021, Fenta et al. (9) conducted a comparative evaluation of six ML models—including logistic regression, LASSO, ridge regression, elastic net, neural networks, and RF across Ethiopian zones and established RF as the top-performing algorithm based on sensitivity, specificity, accuracy, and AUC. That same year, Bitew et al. (10) employed five ML methods, including XGBoost, RF, neural networks, and k-NN to predict socio-demographic risk factors of undernutrition, with XGBoost showing the highest accuracy (88.0%). Khan et al. (11) further validated the utility of ensemble methods by identifying gradient boosting as the most accurate model for predicting stunting among children under five years old. Additionally, Vasu et al. (12) employed the Boruta algorithm in conjunction with RF and PCA for dimensionality reduction to identify the most impactful features in malnutrition prediction. In 2022, Mohammad et al. (13) proposed an optimized hybrid approach combining Harris Hawk Optimization (HHO) with ADASYN for defect prediction, achieving classification accuracies exceeding 99%, with clear implications for imbalanced malnutrition datasets. In 2023, Ndagijimana et al. (14) used ensemble methods including gradient boosting and RF in Rwanda, where gradient boosting achieved the best performance (AUC 89%). The year 2024 has seen even broader adoption: Turjo et al. (15) used six classifiers, including RF and gradient boosting, to assess women’s malnutrition in Bangladesh, where RF showed the highest accuracy and AUC (0.604), Mkungudza et al. (16) applied seven variable selection techniques to logistic regression models to predict undernutrition with modest AUC performance (64%), Boruta feature selection, and the variables’ importance scores were used to identify determinants of malnutrition (17). Yalçın et al. (18) applied elastic net and RF models in neonatal intensive care settings, showing the models’ capacity to accelerate early malnutrition risk identification. This meta-analysis indicated that ML models were observed to have moderate to good performance metrics in predicting malnutrition using DHS data among children under five years (19). Collectively, studies reflect a growing confidence in machine learning models such as SVM, k-NN and ensemble methods, particularly RF. Gradient Boost and XGBoost, for accurate and scalable malnutrition prediction (20–23), though they also reveal ongoing challenges in feature selection, interpretability, and standardization of methodologies.

1.2 Research gap

Machine learning techniques have significantly advanced the identification of malnutrition determinants; they are not without limitations, particularly in how they handle feature importance and inherent model bias. One critical challenge is that statistical variance, feature distribution, and model architecture heavily influence most model-driven feature importance scores (7, 8). As a result, features that may hold substantial contextual or domain relevance can be assigned low importance because they exhibit weak correlations or appear less frequently in the training data. Furthermore, tree-based models like Random Forests or boosting algorithms may exhibit bias towards features with more unique values or features that dominate the data (9–12). This can lead to the underrepresentation of subtle but crucial socio-cultural or

environmental factors. When models are trained on imbalanced or non-representative datasets, they may generalize poorly across regions or demographic groups, further skewing the feature rankings. While multiple studies apply ML algorithms, there is no uniform approach to feature selection, preprocessing, or handling of imbalanced data. Techniques like Boruta, PCA, and ADASYN are used in isolation (12–18), without comparative evaluations or integration into a common pipeline. Most models rely on statistical or wrapper-based feature selection methods (e.g., LASSO, Boruta), but very few studies integrate public health expertise or contextual domain relevance into feature importance ranking (19–21). This leads to the potential exclusion of sociocultural significant predictors that are weakly correlated in raw data but crucial in practice. These challenges make it difficult for policymakers to rely solely on raw model outputs, highlighting the need for more balanced frameworks that integrate statistical learning with human expertise to capture both machine-relevant and domain-important features. To address the limitations of traditional machine learning models in identifying determinants of malnutrition (22, 23), this study proposes the RISE (Relevance-based Integration of Statistics and Expertise) framework. RISE integrates statistical scores, model-based importance, and domain knowledge to uncover both dominant and overlooked features influencing acute forms of child malnutrition.

2 Methods

2.1 Study design, setting, and population

A cross-sectional study was conducted at Mysuru Medical College and Research Center, in collaboration with the Nutrition Rehabilitation Centres (NRC), from March 2024 to January 2025. The sample included 208 children aged 1–23 months admitted to the NRC, and children included in the study were identified with Moderate Acute Malnutrition (MAM) and Severe Acute Malnutrition (SAM), with data on maternal and child anthropometry, and socio-demographic information extracted from hospital records.

2.2 Characteristics of the study population

The study population included child, maternal, and socio-demographic information. The child's information includes 56.8% males and 43.2% females. Children were categorized into four age groups: Group 1 (0–5 months) had 16.0% children, Group 2 (6–11 months) had 40.8% children, Group 3 (12–17 months) included 27.7% children, and Group 4 (18–23 months) comprised 15.5% children. Regarding birth weight, 60.7% children have normal birth weight. Anthropometric measurements showed that the mean Mid-Upper Arm Circumference (MUAC) was 11.57 cm ($\pm$ 2.10 cm), the average weight was 6.32 kg ($\pm$ 1.25 kg), and the mean height was 69.51 cm ($\pm$ 6.71 cm). Among the 206 children included in the study, 51.5% were identified MAM and 48.5% with SAM, reflecting considerable variability in physical growth indicators across the sample. During admission to the NRC, child feeding practices were recorded. The data revealed overlapping patterns; 88.35% of children

received breastfeeding, and 73.79% received bottle feeding. Specifically, 66.02% of children were fed both, 22.33% were exclusively breastfed, 7.77% were exclusively bottle-fed, and 3.88% received neither. These overlaps reflect concurrent feeding practices reported by caregivers and were retained to preserve the full behavioral context in the analysis.

Under maternal characteristics, the majority of mothers, 56.31% in the study were between 20–25 years. 52.91% of the mothers were classified as having an inadequate BMI, highlighting a significant burden of undernutrition among caregivers. Educational attainment with 65.05% having secondary education, though 4.85% remained illiterate. Regarding employment status, a substantial 91.26% were housewives, reflecting low workforce participation. In terms of reproductive profile, 60.19% of mothers had two or more children, and 56.31% of index children were of second or higher birth order, suggesting high family responsibility. 72.82% of mothers had not adopted any family planning methods, indicating a gap in reproductive health awareness and access.

The socio-demographic analysis reveals that 87.86% families reported a non-vegetarian diet. 50.97% of the mothers belonged to Other Backward Classes (OBC), followed by 36.41% from Scheduled Castes and Scheduled Tribes (SC/ST), and 12.62% from the General category. 66.5% of the children were from rural areas. The Ration Card feature in the dataset comprises two categories: Above Poverty Line (APL) and Below Poverty Line (BPL), with BPL accounting for approximately 99.51% of the records. Given this near-constant distribution, a sensitivity analysis was conducted by executing the model both with and without this feature. The results showed no change in model performance metrics or feature rankings. For reference, the model maintained an overall accuracy of 82.54%, as shown in Figure 1, indicating that the inclusion of this low-variance variable did not influence predictive accuracy. This reinforces the robustness of the model and confirms that its outputs are not dependent on statistically redundant features.

2.3 Variables, definitions, and encoding

WHO Anthro software was used to compute Height for Age (HAZ), Weight for Age (WAZ), and Weight for Height (WHZ) scores based on each child's age, height, and weight from NRC records. Nutritional status was classified as SAM (WHZ < −3 SD) and MAM (WHZ between −3 and −2 SD) (24).

Categorical variables in the dataset were numerically encoded to facilitate model training. Binary categorical features such as Breastfeeding Status, Bottle Feeding Status, Family Planning Status, Mother Working Status, Food Type, Child Gender, and Ration card were encoded using label encoding, with values assigned as 1 or 2 to represent distinct categories (e.g., yes/no or male/female). Ordinal variables such as Maternal BMI, Mother MUAC, Child Birth Weight, Maternal Education, and Child Age were label-encoded to preserve their rank structure. Nominal categorical variables like Caste and Residence were similarly encoded using integer labels. All remaining features, including anthropometric and demographic measures such as Child Weight, Child Height, Mother Weight, Mother Height, Child MUAC, and Total Child, were treated as either discrete or continuous numeric variables and

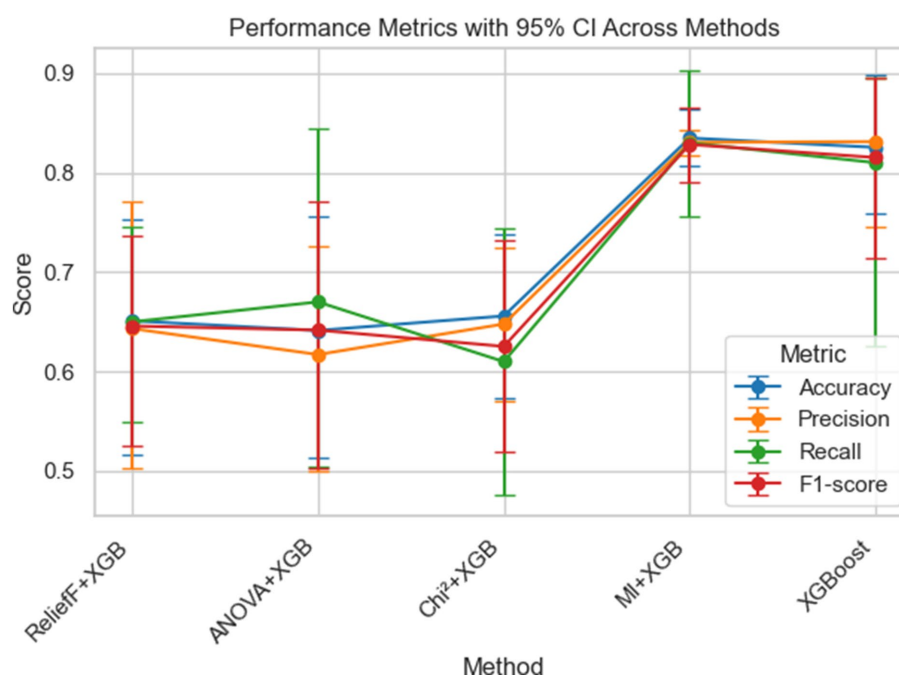


FIGURE 1

Comparative unified performance metrics with 95% confidence intervals across filter-based feature selection methods combined with XGBoost. Comparison of performance metrics with 95% confidence intervals across different filter-based feature selection methods combined with XGBoost. The plot shows Accuracy, Precision, Recall, and F1-score for five methods: Relief + XGB, ANOVA + XGB, Chi² + XGB, MI + XGB, and XGBoost. Bars represent the 95% confidence intervals for each metric.

used without transformation. This encoding strategy ensured compatibility with the modeling pipeline while preserving the interpretability of categorical distinctions.

2.4 RISE framework for feature scoring and selection

The traditional machine learning models rely on feature selection driven primarily by model-based importance scores or a single statistical method such as Mutual Information (MI), Chi-Square (Chi²), or ANOVA for feature scoring. While these approaches can be effective in identifying statistically significant variables, they often place little to no emphasis on domain-driven feature scoring, which can capture contextually important variables that statistical methods may overlook (25–30).

The proposed methodology introduces a novel approach to feature selection through the RISE framework, which is designed to uncover important features that are underrepresented or overlooked by traditional ML models. Conventional model-based scoring methods tend to favor features with strong statistical patterns, frequently ignoring variables that are contextually significant in real-world scenarios. RISE addresses this imbalance by combining four modules:

- (1) A statistical scoring module—captures relevance based on statistical filter methods
- (2) A frequency boost module—provides the boosting scores based on its top priority

- (3) A domain knowledge scoring module—incorporates expert insight and contextual importance
- (4) A model scoring module—reflects importance as assigned by the machine learning model.

This integrated strategy ensures a more balanced and inclusive selection of features in acute forms of child malnutrition domains.

2.4.1 Statistical scoring module

The statistical scoring module in this study leverages filter-based feature selection techniques to evaluate the importance of features based on their statistical relationship with the target variable. The four key methods used are MI, Chi² test, ANOVA F-test, and ReliefF. Each method has its unique way of assessing relevance, which collectively contributes to a more robust feature evaluation.

2.4.1.1 Mutual information (MI)

Mutual Information measures the dependency between a feature and the target variable. It quantifies how much knowing the value of a feature reduces uncertainty about the target. It captures both linear and non-linear relationships. Features with higher MI scores are considered more informative (26, 28).

$$MI(X;Y) = \sum P(x,y) \log \left(\frac{P(x,y)}{P(x) \cdot P(y)} \right)$$

2.4.1.2 Chi-square (Chi²) test

The Chi-Square test assesses whether there is a significant association between a categorical feature and the target class. It

compares the observed frequencies of feature values with the expected frequencies under the assumption of independence. A higher Chi² score indicates a stronger dependency between the feature and the target (28, 29).

$$\chi^2 = \sum \left((O_i - E_i)^2 / E_i \right)$$

2.4.1.3 ANOVA F-test

ANOVA (Analysis of Variance) is used when the feature is continuous and the target is categorical. It evaluates whether the mean of the feature differs significantly across different classes of the target. A high *F*-value suggests that the feature contributes significantly to class separation (27, 28).

$F = MS_B / MS_W$ where $MS_B = SS_B / (k-1)$ and $MS_W = SS_W / (N-k)$

2.4.1.4 ReliefF algorithm

ReliefF is an instance-based feature selection method that considers feature value differences between neighboring instances. It evaluates how well each feature distinguishes between instances of different classes while considering feature interactions and redundancy. It is particularly useful for noisy and complex datasets (30).

$$W[A] = W[A] - (1/m) * \sum (\text{diff}(A, i, \text{nearHit}) - \text{diff}(A, i, \text{nearMiss}))$$

These feature scores, derived before any model training. The selected features based on these scores are then used to train the ensemble boosting model separately for each filter method. The model's internal feature importance scores are subsequently extracted to evaluate how the model prioritizes features that were pre-selected using different statistical criteria. This approach ensures that the model's learning is influenced by features deemed relevant purely through statistical assessment, thereby linking filter-based selection with model-driven evaluation.

2.4.2 Frequency boosting module

The Frequency Boosting Module enhances the weight of features that consistently appear as important across multiple independent selection methods. This approach recognizes that a feature is repeatedly identified by statistical filters based ensemble model. Each time a feature is selected or ranked within the top tier by a model, it earns a frequency point. These counts are then normalized and converted into a frequency boosting score, reflecting the stability and consensus around a feature's importance. By incorporating this module into the RISE framework, features with cross-method agreement are justifiably promoted. This mechanism ensures that consistent signals are amplified, increasing the reliability of the final feature selection.

2.4.3 Domain knowledge scoring module

This Module is designed to integrate expert-driven insights and Contextual relevance into the feature scoring process, leveraging both subject matter expertise and established literature. First, the features

were grouped based on thematic relevance, guided by insights from domain experts and collaborators. These groups represent contextual domains that are well-established in influencing the severity of child malnutrition. Next, each feature was assigned an importance score using XGBoost. The scores were then aggregated within their respective groups to calculate a group-wise cumulative score. Based on these cumulative scores, ranks were assigned to each group. Group 1, Child Anthropometry, includes variables such as gender, age, MUAC, weight, and height of the child (29). Group 2 covers Early Feeding Practices, including birth weight, breastfeeding, and bottle feeding (31, 32). Group 3, Maternal Anthropometry, includes the mother's weight, height, age, BMI, and MUAC (29). Group 4 focuses on Family Structure, including the total number of children and the child's birth order (33). Group 5 includes Socio-Economic Factors such as residence type, ration card status, caste, and food type. Finally, Group 6 represents Maternal Empowerment, including education, working status, and family planning (34). A numerical rank ranging from 6 (highest relevance) to 1 (lowest) was applied at the group level, and each feature within a group inherited the corresponding rank as presented in Table 1. Finally, domain ranks of all features were normalized between 0–1. These normalized scores serve as the final domain importance values in the RISE framework.

TABLE 1 Domain-based feature grouping and prioritization for acute forms of child malnutrition.

Group priority	Group name	Group feature	Group score	Domain rank
1	Child anthropometry	Child gender, age, MUAC, weight, and height	0.275254	6
2	Early feeding practices	Birth weight, breastfeeding, and bottle-feeding practice	0.246018	5
3	Maternal anthropometry	Mother's weight, height, age, BMI, and MUAC	0.147003	4
4	Family structure	Total number of children and the child's birth order	0.127140	3
5	Socio-economic	Residence type, ration card, caste, and food type	0.110287	2
6	Maternal empowerment	Maternal education, working status, and family planning	0.094298	1

Domain-wise grouping and ranking of features influencing acute forms of child malnutrition. Each group consists of thematically related variables aggregated to calculate a Group Score, which indicates the combined influence of that domain. Domain Rank shows the relative priority based on these scores, with higher values indicating greater importance.

2.4.4 A model scoring module

The model scoring module independently generates feature importance scores directly from a boosting model without relying on any prior feature selection or pre-training. In this module, the full set of input features is used to train an ensemble boosting algorithm where the model itself learns which features contribute most to predictive performance.

2.4.5 Final feature scores

The RISE framework computes a final feature score by integrating model-based performance, methodological consistency, and domain relevance. This scoring strategy ensures that selected features are not only powerful predictors but also consistently recognized across methods and grounded in expert knowledge.

The Model Score reflects the feature’s contribution to prediction accuracy, derived directly from a boosting model without any prior filtering. The Frequency boost captures how often a feature is selected as important across a filter-based ensemble model, indicating its robustness and consistency. The Domain Importance represents the contextual and theoretical significance of a feature as determined by thematic grouping guided by subject matter experts.

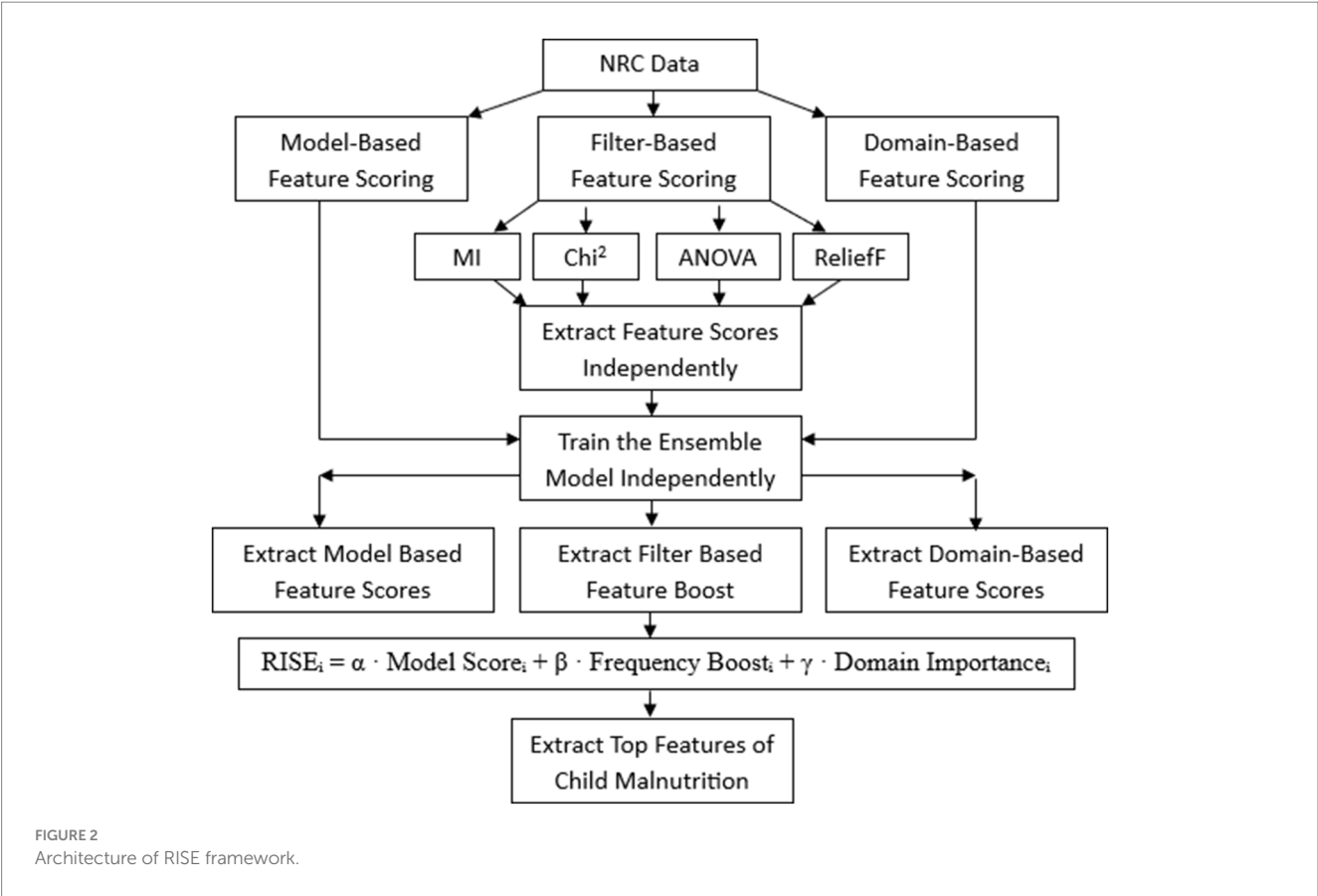
RISE Final Score Formula:

$$RISE_i = \alpha \cdot \text{Model Score}_i + \beta \cdot \text{Domain Importance}_i + \gamma \cdot \text{Frequency Boost}_i$$

i : Feature index. α, β, γ : Hyper-tuning values to adjust the contribution of each component. In the RISE framework, Selection of Hyperparameter tuning was done using nested validation. The final RISE score is defined as a weighted sum of Model Score, Domain importance, and Frequency Boost. The optimal configuration was determined to be $\alpha = \beta = \gamma = 0.1$. The overall architecture of the RISE framework is presented in Figure 2.

3 Results

The dataset used in this study comprises exclusively malnourished children, with the target variable categorically divided into SAM and MAM. The primary objective of this research is to identify the determinants of acute forms of child malnutrition. The analysis focuses on extracting and ranking features through a combination of model-based importance, domain knowledge insights, and frequency-based boosting grounded in a filter-based method. *A priori* power analysis using G*Power (two-tailed, effect size $d = 0.5$, $\alpha = 0.05$, power = 0.90) indicated that a minimum of 172 participants (86 per group) would be required. This study includes real-time data of 206 children (100 SAM and 106 MAM), ensuring adequate statistical power as well as indicating a balanced distribution between the two groups. The dataset was obtained from the NRC register. All available features were taken for analysis to ensure a comprehensive evaluation and to support the development of a framework for identifying



determinants of acute forms of child malnutrition. The study did not use any confounder study, pre-filtering, or exclusion based on correlation thresholds. It retained the full spectrum of variables that may contribute to nutritional outcomes of the child, whether directly or indirectly. This inclusive approach allows the model to assess the relative importance of each feature in context, which ensures completeness and eliminates concerns regarding missing values.

The implementation was done in Python version 3.11.11 (Anaconda distribution). The dataset was first partitioned into training and testing subsets to evaluate generalization performance. Feature selection was guided by XGBoost classifier, statistical importance based on filter-based XGBoost, and domain ranking with XGBoost. To ensure robustness and reduce overfitting, Stratified K-Fold Cross-Validation (with $k = 5$) was employed on the training set. This approach preserved the class distribution across folds and allowed for consistent performance estimation. A nested cross-validation framework was employed for hyperparameter tuning. The outer loop used 5-fold cross-validation to evaluate generalization performance, while the inner loop applied 3-fold cross-validation within a grid search. The optimal values obtained were $\alpha = \beta = \gamma = 0.1$, resulting in a Nested CV mean F1 score of 0.824. For the XGBoost classifier, parameters specified were the number of estimators ($n_{\text{estimators}} = 100$), evaluation metric ($\text{eval_metric} = \text{logloss}$), and random seed ($\text{random_state} = 42$). All other hyperparameters were retained at their default values ($\text{learning_rate} = 0.1$, $\text{max_depth} = 6$, $\text{subsample} = 1.0$).

The XGBoost, a high-performance gradient boosting algorithm, is widely used for its accuracy, speed, and ability to handle complex, structured data (35–38). This section presents the experimental implementation of the RISE framework using the XGBoost model. However, the proposed approach is flexible and can be generalized to other machine learning ensemble methods for enhanced applicability. The results of model-based and domain-based feature scoring are summarized in Table 2. The model-based scores were derived by training an XGBoost classifier independently on a curated set of input features and extracting importance, which reflects each feature's contribution to child malnutrition. In parallel, domain ranks were obtained based on thematic relevance, guided by insights from domain experts and group score, and it is further normalized to ensure symmetry in the RISE score components, as presented in Table 1.

To identify the most influential factors associated with acute forms of child malnutrition, filter-based feature selection methods were employed. Each method prioritized the top 15 features based on their statistical relevance to the target outcome. These selected features were used independently to train XGBoost models. This approach allowed us to assess how feature selection strategies influence model performance and feature importance scoring. Further, each feature earns its frequency boost score based on the number of times it appears in the top 10 ranked features across four different filter-based feature selection methods. If a feature appears in the top 10 of all four methods, it is assigned a boost value of 4. If it appears in three methods, the boost value is 3; in two methods, it is 2, and in one method, it is 1. If the feature does not appear in the top 10 of any method, the frequency boost is 0. This scoring mechanism reflects the consistency and recurrence of feature importance across multiple selection methods. The results of filter-based XGBoost, along with frequency boost, are presented in Table 3. The performance of the

TABLE 2 Feature importance from XGBoost model and normalized domain scores.

Feature	Model importance	Domain rank (normalized)
Child weight	0.1304	1.0
Child MUAC	0.0989	1.0
Breastfeeding status	0.0969	0.8
Child height	0.0820	1.0
Child order	0.0669	0.4
Mother age	0.0550	0.6
Food type	0.0539	0.2
Mother weight	0.0485	0.6
Child gender	0.0453	1.0
Maternal BMI	0.0443	0.6
Mother height	0.0443	0.6
Mother MUAC	0.0411	0.6
Residence	0.0340	0.2
Maternal education	0.0336	0.1
Child birth weight	0.0279	0.8
Total child	0.0252	0.4
Child age	0.0248	1.0
Bottle feeding status	0.0220	0.8
Caste	0.0139	0.2
Family planning status	0.0112	0.1
Maternal WORKING STATUS	0.0000	0.1
Ration card	0.0000	0.2

Presents XGBoost model importance and normalized domain scores. The domain scores are adapted from domain-derived weightings as shown in Table 1 and further normalized between 0–1.

XGBoost model across various feature training configurations, its accuracy, recall, precision, and F1-scores are presented as a unified performance report in Figure 1.

Now, the RISE framework is employed by integrating three key components: domain importance scoring, model-based feature scoring, and frequency-based boosting. Each feature's final score is computed by combining these elements using a weighted formula, enabling a balanced representation of statistical relevance, expert knowledge, and selection consistency presented in Table 4. This comprehensive scoring strategy ensures that both high-performing and domain-relevant but underrepresented features are prioritized appropriately in the final analysis, visualized in Figure 3.

4 Discussion

The analysis of top-ranked features using the RISE framework reveals critical insights into the multifaceted nature of acute forms of child malnutrition. Among the various determinants, child anthropometric indicators emerged as the most influential, followed by maternal anthropometry and child order. However, Data from the

TABLE 3 Consolidated feature importance scores from XGBoost models using multiple filter methods and associated frequency boosting.

SL. no	Feature	MI+XGB	CHI2 + XGB	ANOVA+XGB	ReliefF + XGB	Frequency boost
1	Child weight	0.17623	0.07254	0.07823	0.08677	4
2	Mother weight	0.07228	0.06940	0.06450	0.08097	4
3	Child age	0.05977	0.05965	0.04656	0.07366	4
4	Child order	0.11333	0.08627	0.07579	0.09466	4
5	Mother MUAC	0.03908	0.08464	0.06795	0.10448	3
6	Maternal BMI	0.07584	0.03560	0.04508	0.07273	3
7	Maternal education	0.02548	0.07033	0.10025	0.13661	3
8	Food type	-	0.22474	0.30430	-	2
9	Child MUAC	-	-	0.05691	0.08130	2
10	Child gender	0.09670	0.06540	-	-	2
11	Family planning status	0.06051	-	-	0.05844	2
12	Child birth weight	0.04554	0.05052	0.04102	0.04506	2
13	Residence	-	0.05299	-	-	1
14	Bottle feeding status	0.07320	-	-	-	1
15	Total child	0.01176	-	-	0.06862	1
16	Mother age	0.01176	0.05808	0.02073	-	1
17	Child height	0.13852	-	-	-	1
18	Mother height	-	-	0.03518	0.04935	0
19	Breastfeeding status	-	-	-	0.04736	0
20	Caste	-	0.04637	0.04038	-	0
21	Maternal working status	-	0.02346	0.02313	0.000000	0
22	Ration card	0.000000	0.000000	0.000000	0.000000	0

The table presents consolidated feature importance scores for acute forms of child malnutrition. The importance scores are obtained from XGBoost models trained using various filter methods: MI, CHI2, ANOVA, and ReliefF. Only features that appear in the top 10 of at least one configuration were given frequency boost score. The Frequency Boost Score indicates the feature's presence across top-10 lists: 4 – feature appears in the top 10 of all 4 methods. 3 – feature appears in the top 10 of any 3 methods. 2 – feature appears in the top 10 of any 2 methods. 1 – feature appears in the top 10 of any 1 method. 0 – feature does not appear in the top 10 of any method.

NRC reinforces these findings: over 88% of malnourished children were being breastfed, including 58.51% of MAM cases and 44.32% of SAM cases. Furthermore, 49.0% of mothers of children diagnosed with SAM had an inadequate BMI, indicating undernutrition. Similarly, a higher proportion, 56.6% of mothers of children with MAM, also exhibited inadequate BMI levels (8). This strongly suggests an interlinked pattern that mothers with poor nutritional status may be breastfeeding children who consequently face an increased risk of malnutrition, emphasizing the dual burden of maternal and child undernutrition. The study also implements SHAP (SHapley Additive Explanations), which indicates how each feature contributes to the model's predictions. The SHAP results further validate that maternal anthropometry, such as the mother's weight, plays a significant role in acute forms of child malnutrition presented in Figure 4. Child order reflects the burden of familial resource allocation, with 56.31% of malnourished children being of second birth order or higher, indicating that increasing family size may dilute maternal attention and care (9, 10), further confirmed with SHAP analysis presented in Figure 4. Together, the integrated model insights and contextual statistics highlight critical intervention points. The evidence highlights the importance of policy strategies that enhance maternal nutrition,

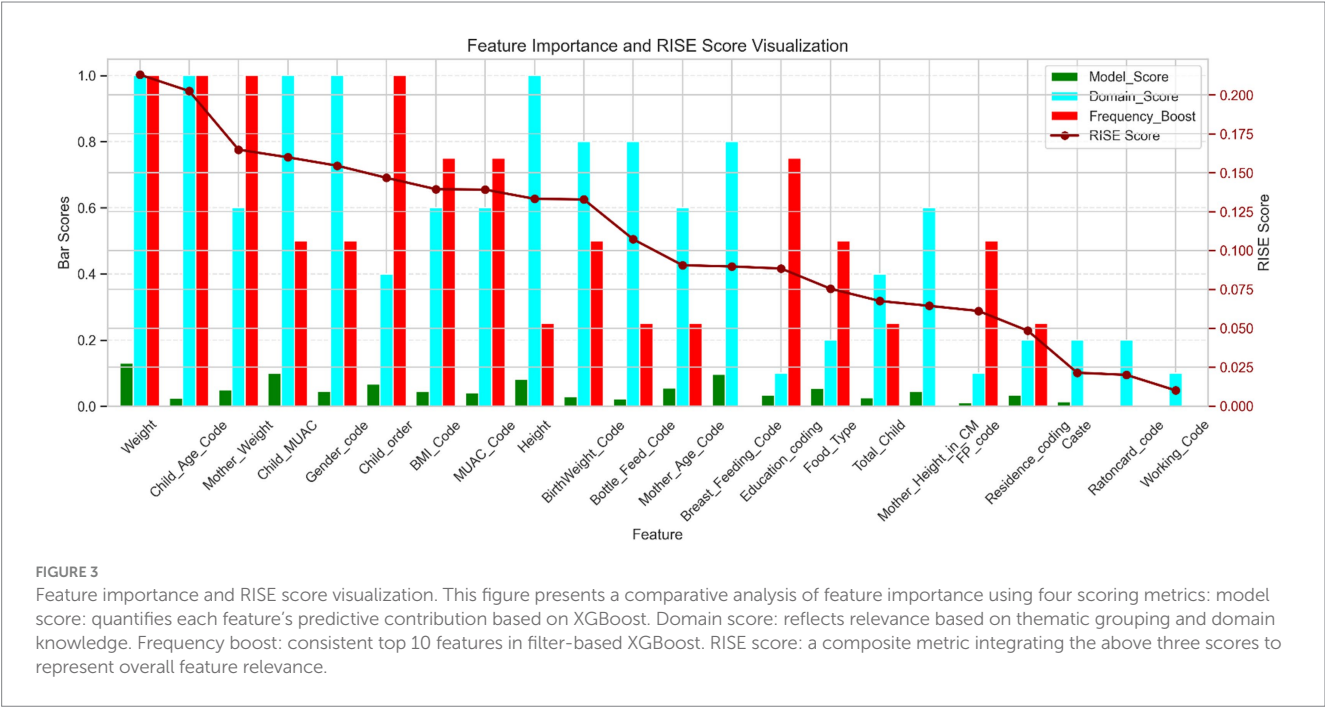
target vulnerable groups, and support young mothers, particularly those breastfeeding, in effectively breaking the intergenerational cycle of malnutrition (29, 30).

This research offers a novel contribution to the field of Child malnutrition and Machine learning. Firstly, in the child malnutrition domain, multi-dimensional and data-driven exploration of the complex determinants of acute forms of child undernutrition is achieved through integrated analysis of Maternal and Child Factors, emphasizing the mother–child nutritional dyad as a central axis of malnutrition (29). This integrated approach moves beyond child-focused indicators and reflects a life-cycle and intergenerational perspective on malnutrition. Secondly, this research introduces a novel and comprehensive feature prioritization framework, RISE (Relevance-based Integration of Statistics and Expertise), that represents a significant advancement in the field of child malnutrition analytics. Unlike conventional studies that depend on machine learning feature importance or subjective domain prioritization (7–17), this study uniquely integrates multi-dimensional evidence by combining four robust statistical filter methods (Mutual Information, Chi-square, ANOVA-F, ReliefF) with an ensemble model, domain-based scoring, and a newly proposed frequency boosting technique

TABLE 4 Final RISE with normalized domain score, frequency boost, and domain score with corresponding hyper tuning values.

Feature	Domain score	Model score	Frequency boost	RISE score
Child weight	1.0	0.1304	1.00	0.2130
Child age	1.0	0.0248	1.00	0.2025
Mother weight	0.6	0.0485	1.00	0.1648
Child MUAC	1.0	0.0989	0.50	0.1599
Child gender	1.0	0.0453	0.50	0.1545
Child order	0.4	0.0669	1.00	0.1467
Maternal BMI	0.6	0.0443	0.75	0.1394
Mother MUAC	0.6	0.0411	0.75	0.1391
Child height	1.0	0.0820	0.25	0.1332
Child birth weight	0.8	0.0279	0.50	0.1328
Bottle feeding status	0.8	0.0220	0.25	0.1072
Mother age	0.6	0.0550	0.25	0.0905
Breastfeeding status	0.8	0.0969	0.00	0.0897
Maternal education	0.1	0.0336	0.75	0.0884
Food type	0.2	0.0539	0.50	0.0754
Total child	0.4	0.0252	0.25	0.0675
Mother height	0.6	0.0443	0.00	0.0644
Family planning status	0.1	0.0112	0.50	0.0611
Residence	0.2	0.0340	0.25	0.0484
Caste	0.2	0.0139	0.00	0.0214
Ration card	0.2	0.0000	0.00	0.0200
Working code	0.1	0.0000	0.00	0.0100

The table presents consolidated feature scores for malnutrition-related data. The RISE_Score represents the final integrated score for each feature, calculated by combining: 1. Domain score – provided based on thematic grouping and expert ranking and further normalized between 0–1. 2. Model score – provides the feature importance obtained from the XGBoost classifier. 3. Frequency boost – provided for the feature that is repeatedly identified by statistical filters based ensemble model. The combination is controlled by hyperparameters α , β , and γ , which were optimized using nested cross-validation (nested CV). The optimal values obtained were $\alpha = \beta = \gamma = 0.1$, resulting in a nested CV mean F1 score of 0.824.



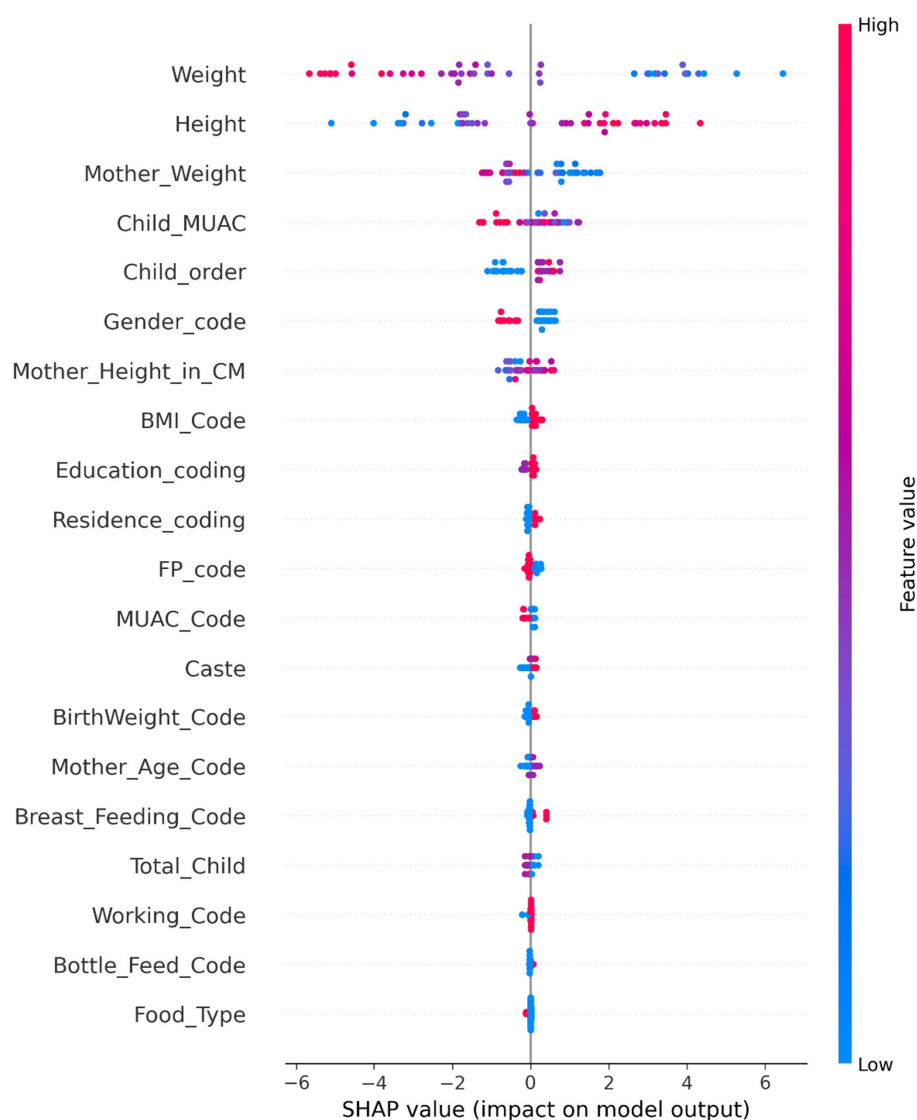


FIGURE 4
SHAP summary plot showing feature contributions for acute forms of child malnutrition.

that captures the cross-method consensus of feature relevance. This multi-branch integration ensures that underrepresented yet contextually vital variables are not bypassed by purely statistical weightages (27). The RISE framework thus reflects a paradigm shift from purely model-driven selection to an interpretable, explainable, and policy-relevant decision layer. It is an adaptable scoring logic and modular design, suitable for extension to other health and development research areas, establishing both scientific novelty and practical utility.

Some features, such as Mother Height, Breastfeeding Status, Caste, Maternal Working Status, and Ration Card, obtained a frequency boost score of zero. This indicates that none of the statistical feature selection methods consistently identified these variables as important predictors. The likely reasons include weak or indirect associations with acute forms of malnutrition and underrepresentation of certain subgroups in the dataset. Nevertheless, these features remain highly relevant from a public health and domain perspective, as they capture socioeconomic and

behavioral dimensions of child nutrition that purely statistical methods may overlook. Therefore, the RISE framework incorporates domain-based correction to ensure such features are not disregarded solely based on low statistical detectability. Limitations of this study include the potential lack of generalizability due to the region-specific and institutionalized nature of the dataset. Such data may reflect only children accessing a particular facility, thereby excluding those without access to the facility or in different regional contexts. Additionally, while the RISE framework offers an advantage by incorporating domain knowledge along with statistical and model-driven methods, it introduces a level of subjectivity, particularly in assigning domain-based weights. Expert input for feature grouping and ranking was not derived through a formal elicitation process. External and temporal validation are essential to ensure the model's reliability across different populations, settings, and time periods. In the study, the model demonstrates strong performance within the current dataset; its generalizability remains untested.

5 Conclusion

By examining the top features ranked by the RISE Score, Child anthropometry emerged as the most influential, followed by maternal anthropometry and child order. This hierarchy underscores a double burden of malnutrition. This reflects the importance of physical growth parameters in assessing nutritional status. Further down, features like maternal anthropometry and child order underscore its influence on child health outcomes. The presence of these variables at the top reinforces the understanding that the importance of maternal nutritional status and physical attributes plays a key role in shaping a child's growth and development, further highlighting the multifactorial nature of malnutrition.

The RISE framework demonstrates its strength by effectively identifying and assigning importance to features that were under-prioritized by conventional filter-based or model-based feature selection methods. For instance, features such as Ration card = 0.0200, Working status = 0.0100, Caste = 0.0100, Mother height = 0.0644, and Breastfeeding status = 0.0897 lacked frequency boost contributions but were still given significant weight by the RISE framework. This reflects the inclusive nature of RISE, which integrates domain relevance, statistical contribution, and contextual importance, allowing it to highlight features that might be overlooked by purely data-driven models but are crucial in real-world health and nutritional contexts.

Addressing child malnutrition is a crucial step toward achieving Sustainable Development Goals, SDG 2 (Zero Hunger) and SDG 3 (Good Health and Well-being). The development and implementation of the RISE framework is an advancement in malnutrition analytics by bridging the gap between data-driven methodologies and domain expertise. Unlike traditional feature selection methods that may overlook critical variables due to statistical bias, RISE facilitates a more equitable, context-aware, and interpretable feature evaluation. By elevating the significance of factors like maternal education, breastfeeding practices, and socioeconomic conditions, the RISE approach ensures that key public health insights are retained and highlighted in machine learning models. Finally, this framework strengthens the translation of empirical data into actionable knowledge, empowering targeted interventions and evidence-based policymaking in the process of child malnutrition, thus contributing directly to global health equity and the realization of the SDGs.

5.1 Future work

While the current model demonstrates strong performance within the study dataset, to ensure reliability, broader applicability, or generalizability, future research should have a strong focus on external and temporal validation. Validating the framework on independent datasets from different geographic regions and across varied time periods, such validation would help in identifying the key determinants of acute child malnutrition.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed

to the corresponding author. The code is available at: <https://github.com/Shruthi-S-Scholar/ChildMalnutrition-rise-framework>.

Ethics statement

The studies involving humans were approved by Mysore Medical College and Research Institute, Mysore, Karnataka. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

SShr: Writing – original draft, Writing – review & editing. PG: Conceptualization, Supervision, Validation, Writing – review & editing. SSha: Data curation, Supervision, Validation, Writing – review & editing. PA: Supervision, Writing – review & editing. AU: Supervision, Writing – review & editing. LR: Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. World Health Organization. (2024). Malnutrition. Available online at: <https://www.who.int/news-room/fact-sheets/detail/malnutrition>.
2. World Health Organization, UNICEF, & World Bank Group. (2025). Levels and trends in child malnutrition: UNICEF/WHO/World Bank Group joint child malnutrition estimates—Key findings of the 2025 edition (ISBN 978–92–4–011230–8). Available online at: <https://www.who.int/publications/i/item/9789240112308>.
3. IIPS & MoHFW. (2021). International Institute for Population Sciences (IIPS) & Ministry of Health and Family Welfare (MoHFW). Ministry of Health and Family Welfare, India. Available online at: https://mohfw.gov.in/sites/default/files/NFHS-5_Phase-II_0.pdf
4. Laksono AD, Izza N, Trisnani T, Paramita A, Sholikhah HH, Andarwati P, et al. Determination of appropriate policy targets to reduce the prevalence of stunting in children under five years of age in urban-poor communities in Indonesia: a secondary data analysis of the 2022 Indonesian national nutritional status survey. *BMJ Open*. (2024) 14:e089531. doi: 10.1136/bmjopen-2024-089531
5. Rezaee MT, Zakki SA, Haq I u, Rahimi N, Fayaz M. Cross-sectional study of determinants of undernutrition among children aged 6–36 months in Kabul, Afghanistan. *BMJ Open*. (2024) 14:e079839. doi: 10.1136/bmjopen-2023-079839
6. Qasrawi R, Sgahir S, Nemer M, Halaikah M, Badrasawi M, Amro M, et al. Machine learning approach for predicting the impact of food insecurity on nutrient consumption and malnutrition in children aged 6 months to 5 years. *Children*. (2024) 11:7. doi: 10.3390/children11070810
7. Anku EK, Duah HO. Predicting and identifying factors associated with undernutrition among children under five years in Ghana using machine learning algorithms. *PLoS One*. (2024) 19:e0296625. doi: 10.1371/journal.pone.0296625
8. Talukder A, Ahammed B. Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. *Nutrition*. (2020) 78:110861. doi: 10.1016/j.nut.2020.110861
9. Fenta HM, Zewotir T, Muluneh EK. A machine learning classifier approach for identifying the determinants of under-five child undernutrition in Ethiopian administrative zones. *BMC Med Inform Decis Mak*. (2021) 21:291. doi: 10.1186/s12911-021-01652-1
10. Bitew FH, Sparks CS, Nyarko SH. Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia. *Public Health Nutr*. (2022) 25:269–80. doi: 10.1017/S1368980021004262
11. Khan JR, Tomal JH, Raheem E. Model and variable selection using machine learning methods with applications to childhood stunting in Bangladesh. *Inform Health Soc Care*. (2021) 46:425–42. doi: 10.1080/17538157.2021.1904938
12. Vasu SR, Khare S, Gupta D, Jyotishi A. Features explaining malnutrition in India: a machine learning approach to demographic and health survey data In: D Garg, K Wong, J Sarangapani and SK Gupta, editors. Advanced computing. 10th International Conference, IACC 2020, Panaji, Goa, India: Springer (2021). 87–99.
13. Mohammad UG, Imtiaz S, Shakya M, Almadhor A, Anwar F. An optimized feature selection method using ensemble classifiers in software defect prediction for healthcare systems. *Wirel Commun Mob Comput*. (2022) 2022:1028175. doi: 10.1155/2022/1028175
14. Ndagijimana S, Kabano IH, Masabo E, Ntaganda JM. Prediction of stunting among Under-5 children in Rwanda using machine learning techniques. *J Prev Med Public Health*. (2023) 56:41–9. doi: 10.3961/jpmph.22.388
15. Turjo EA, Rahman MH. Assessing risk factors for malnutrition among women in Bangladesh and forecasting malnutrition using machine learning approaches. *BMC Nutr*. (2024) 10:22. doi: 10.1186/s40795-023-00808-8
16. Mkungudza J, Twabi HS, Manda SOM. Development of a diagnostic predictive model for determining child stunting in Malawi: a comparative analysis of variable selection approaches. *BMC Med Res Methodol*. (2024) 24:175. doi: 10.1186/s12874-024-02283-6
17. Saleem J, Zakar R, Butt MS, Aadil RM, Ali Z, Bukhari GMJ, et al. Application of the Boruta algorithm to assess the multidimensional determinants of malnutrition among children under five years living in southern Punjab, Pakistan. *BMC Public Health*. (2024) 24:167. doi: 10.1186/s12889-024-17701-z
18. Yalçın N, Kaşıkçı M, Çelik HT, Demirkan K, Yiğit Ş, Yurdakök M. Development and validation of machine learning-based clinical decision support tool for identifying malnutrition in NICU patients. *Sci Rep*. (2023) 13:5227. doi: 10.1038/s41598-023-32570-z
19. 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) 22:449. doi: 10.3390/ijerph22030449
20. Zheng F, Chen K, Zhang X, Wang Q, Zhang Z, Niu W. Prevalence of malnutrition and associated factors in Chinese children and adolescents aged 3–14 years using machine learning algorithms. *J Glob Health*. (2025) 15:04204. doi: 10.7189/jogh.15.04204
21. Ayele MK, Baye GA, Yesuf SH, Engda AA, Mitiku ET. Predicting stunting status among under five children in Ethiopia using ensemble machine learning algorithms. *Sci Rep*. (2025) 15:27907. doi: 10.1038/s41598-025-03206-1
22. 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. doi: 10.1016/j.nut.2024.112631
23. Sugihartono T, Wijaya B, Marini M, Alkayess AP, Anugerah HA. Optimizing stunting detection through SMOTE and machine learning: a comparative study of XGBoost, random forest, SVM, and k-NN. *J Appl Data Sci*. (2025) 6:667–82. doi: 10.47738/jads.v6i1.494
24. World Health Organization (WHO) (2006). Child growth standards. Available online at: <https://www.who.int/publications/i/item/924154693X>
25. Paul S, Rahman M, Dolley A, Saikia K, Singh CS, Mohammed A, et al. A retrospective study using machine learning to develop predictive model to identify rotavirus-associated acute gastroenteritis in children. *PeerJ*. (2025) 13:e19025. doi: 10.7717/peerj.19025
26. Tripathi A, Ragiri PR, Jain D, Yadav T. Machine learning-based predictive models for early diagnosis of liver disease. *J Sci Ind Res*. (2025) 84, 575–583. doi: 10.56042/jsir.v84i5.14828
27. Rabie AH, Aldawsari M, Saleh AI, Saraya MS, Rashad M. HFSA: hybrid feature selection approach to improve medical diagnostic system. *PeerJ Comput Sci*. (2025) 11:e2764. doi: 10.7717/peerj-cs.2764
28. Soliman MM, Chowdhury MH, Murugappan M, Chowdhury MEH. Automated classification of post-operative gait abnormalities following hip surgery using machine learning. *Eng Res Exp*. (2025) 7:035203. doi: 10.1088/2631-8695/ade84d
29. Mekonnen S, Birhanu D, Menber Y, Gebreegziabher ZA, Belay MA. Double burden of malnutrition and associated factors among mother–child pairs at household level in Bahir Dar City, Northwest Ethiopia: community based cross-sectional study design. *Front Nutr*. (2024) 11:1340382. doi: 10.3389/fnut.2024.1340382
30. Rabiei S, Ebrahimof S, Rasekhi H, Amini M, Ghodsi D, Yari Z, et al. Exploring the determinants of malnutrition in 2–5 year Iranian children using structural equation modeling: national food and nutrition surveillance. *BMC Public Health*. (2024) 24:3406. doi: 10.1186/s12889-024-20931-w
31. Kofi Amegah A, Ayinemi R, Sewor C, Fenta HM, Yeboah K, Mohammed SA, et al. Birth weight mediates the association of maternal undernutrition with child undernutrition prevalence in West Africa. *Eur J Clin Nutr*. (2024) 78:772–81. doi: 10.1038/s41430-024-01453-5
32. Chiopris G, Chiopris C, Valenti M, Esposito S. Determinants of undernutrition among children admitted to a Pediatric Hospital in Port Sudan, Sudan. *Nutrients*. (2024) 16:787. doi: 10.3390/nu16060787
33. Riwa FP, Odgers-Jewell K, Jones MA, Mushi AA. The prevalence and determinants of undernutrition among infants and children aged 6 months to 5 years in sub-Saharan African countries: a systematic scoping review. *Nutr Rev*. (2025) 83:e1896–916. doi: 10.1093/nutrit/nuae189
34. Pandey S, Rahut DB, Araki T. Ethnicity/caste and child anthropometric outcomes in India using the National Family Health Survey 2015–16 and 2019–21. *PLoS One*. (2024) 19:e0311092. doi: 10.1371/journal.pone.0311092
35. Bardia V, Sophiya E.. (2024). Diabetes prediction using machine learning algorithm: a comparative analysis. 2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS), 1973–1979.
36. Km KR, Khan SB, Govindarajan P, TR M, Alojail M, Gadekallu TR. Machine learning-driven intelligent water quality assessment for enhanced drinking safety and real-time consumer awareness. *Hydrol Res*. (2025) 56:136–52. doi: 10.2166/nh.2025.097
37. Bonthu Y., Mannam S., Kandikunta G., Keshagani V. G., Sarath G. (2024). Heart attack risk prediction using advanced machine learning techniques. 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), 1–6.
38. Reddy J. V. N. Y., Ram G. S. S., Varshini V., Srinivas M.. (2025). An empirical evaluation of machine learning models for stroke risk assessment with explainable AI. 2025 International Conference on Artificial Intelligence and Data Engineering (AIDE), 150–157.