International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Novel Adaptive Feature-Penalized Ridge Logistic Regression Model: A Supervised Machine Learning Advancements Over Traditional Ridge Logistic Regression for Composite Malnutrition Diagnosis Among Under-Five Children in Mozambique

Adelino Benedito Nhancale[1], Fredrick Onyango[2], Joyce Otieno[3]

[1]Lecturer, Faculty of Economics, Eduardo Mondlane University
[2]Lecturer, Department of Statistics & Actuarial Science, Maseno University
[3]Lecturer and Head of the Department, Department of Statistics & Actuarial Science, Maseno University

Abstract

In Mozambique, child malnutrition is not only a pressing public health issue but also one of the most complex, multi-faceted, persistent problems, where stunting, underweight, and wasting continue to defy numerous efforts. Efforts to understand and tackle the issues of malnutrition have been hindered by traditional statistical techniques such as logistic regression due to multicollinearity and interdependence of these outcomes situated within an intricate causal framework. We present a new statistical technique: Adaptive Feature-Penalized Ridge Logistic Regression (AFPR-RM), which models a composite malnutrition outcome by capturing stunting (chronic), underweight (intermediate), and wasting (acute) into a singular term for children below 5 years of age. Using the 2022 Mozambique Demographic and Health Survey (DHS), we extract 3,953 observations. We present Ridge Logistic Regression (RLR) model which overcomes the inadequacy demonstrated by classical logistic regression via multicollinearity. AFPR-RM is advanced by incorporating a dual-adaptive penalty structure: one driven by the data-derived signal strength and the other, feature scaling determined from domain-specific considerations. Together, these create a multiplicative penalty framework, which we deem a novel contribution to the literature on penalised regression. AFPR-RM outperforms standard Ridge Logistic Regression in classification accuracy, model interpretability, and diagnostic stability, as shown by empirical comparison. This study aims to assess the classification, accuracy, calibration and diagnostics of the AFPR-RM against the baseline Ridge Logistic Regression model and determine the best model for guiding public health interventions on child malnutrition across its indicators.

Keywords: Child malnutrition, DHS, Mozambique, Multicollinearity, Ridge Logistic Regression (RLR), Adaptive Feature-Penalised Ridge Logistic Regression (AFPR-RM).

1. Introduction

Global public health challenges still exist in bearing the burden of malnourished children under the age of five. It significantly impairs their growth, cognitive development, and health in general. In sub-Saharan Africa, there is still a high prevalence of malnutrition in the form of stunting, wasting, and underweight even after several attempts to intervene Black et al. (2008), UNICEF (2023). In Mozambique, attempts to reduce malnutrition have been frustrated by intricate socio-economic, infrastructural, and maternal health factors, Instituto Nacional de Estatística & ICF (2024).

In addressing malnutrition and formulating targeted policies for intervention, statistical modelling becomes very important from a policy perspective. One of the regularisation techniques that is widely used is Ridge Logistic Regression (RLR), which deals with multicollinearity by shrinking regression coefficients, Nhancale et al. (2025), Hoerl & Kennard (1970). However, RLR's uniform penalization may overlook nuances in feature importance. Adaptive penalized models such as the Adaptive Feature-Penalized Ridge Regression Model (AFPR-RM) introduced in this paper, apply variable-specific penalties allowing for greater precision and reliable inference within large complex health datasets, Zou (2006), Hastie et al. (2009).

Recent studies have pointed out the problems associated with standard regression methods dealing with collinearity and composite health outcomes interpretability, Tibshirani (1996), Vounatsou et al. (2019). Moreover, Ridge models have been used in public health, as Chui MK et al. (2024) noted, but their effectiveness in comparison to adaptive models is still under-studied, especially in African contexts.

Using Mozambique's DHS 2022 dataset, this study aims to analyse and compare the predictive and inferential capabilities of RLR and AFPR-RM. It looks at model specific performances in the identification of key malnutrition determinants, analyses classification accuracy, and investigates the interpretative worth of adaptive penalisation. Qualitatively, the paper aims to be a contribution towards the methodological enhancement of malnutrition modelling in resource-poor settings through the proposed comparison approach.

2. Literature Review

Among the many forms of health-related issues that require urgent attention, childhood malnutrition is one of the most prominent, especially in low- and middle-income countries. Black et al. (2008) provide evidence regarding the burden of undernutrition of mothers and children and its effects on child mortality and developmental deficits later in life. The UNICEF (2023) Report added that the COVID-19 pandemic has halted a decade's worth of progress towards immunisation and nutrition for children, which adds to the problem at hand. De Onis and Branca (2016) pointed out that stunting remains a vital form of child malnutrition with its prevalence rate standing at 161 million children globally.

Older models for child malnutrition used to focus on determinants such as household wealth, maternal education, and birthweight using logistic regression. Fotso (2007) studied child malnutrition, focusing on the urban-rural differentials. Some of the findings from this study looked at disparities that were associated with socio-economic factors. Kandala et al. (2011) undertook a scoping review on the factors associated with malnutrition in children under five years old in Sub-Saharan Africa. In another study on assessing factors associated with multifactorial determinants of under-five child malnutrition in Mozambique, as an alternative to the classical logistic regression due to multicollinearity factor in DHS data, Nhancale et al. (2025) used Ridge Logistic Regression model to predict a binary composite malnutrition outcome, having concluded that Ridge Logistic Regression is efficient in controlling multicollinearity as well as allowing consistent estimation over many dependent interrelated predictors. Some of the important factors noted were maternal education and household wealth. However, these traditional methods using Demographic and Health Surveys (DHS) datasets tend to face multicollinearity problems and overfitting of the data. To overcome these problems, penalised regression techniques have been utilised.

First discussed in 1970 by Hoerl and Kennard, ridge regression utilises L2 penalisation to mitigate issues of multicollinearity when estimating coefficients. In 2009, Hastie and colleagues built upon this concept by discussing various techniques of regularisation in statistical learning. Zou (2006) presented the adaptive lasso which selectively imposes greater penalties to improve variable selection and increase precision of the model. Friedman et al. (2010) provided efficient algorithms for computing regularisation paths for generalised linear models which enabled the use of such techniques in large scale datasets.

Within the domain of child health, these models have proven useful. In 2019, Kim and colleagues, through the multi-variable modelling approach, were able to determine child health outcome predictors, indicating the potential of these methods in complex epidemiological data. Gebresilassie et al. (2020), applying ridge regression, assessed predictors of anaemia in Ethiopian children and demonstrated the method's strength in dealing with multicollinearity. Tessema et al. (2021) used adaptive lasso techniques to examine the factors associated with neonatal mortality, reinforcing the notion that penalised regression is useful in health research.

However, very few studies have utilized adaptive ridge types, like the Adaptive Ridge Regression Model, in relation to composite malnutrition outcomes. Myatt et al. (2018) supported the use of composite indicators in child nutrition research because they believed that disaggregating stunting, wasting, and underweight was a poor attempt at capturing the complexity of malnutrition. Grellety and Golden (2018) argued that better classification systems are needed to accurately define and treat severe acute malnutrition. To model such composite indicators requires specialised techniques and methods that can manage and intertwine intricate dependencies and relations among numerous variables, which is a gap where adaptive penalised models shine, Van de Wiel et al., (2021).

In the context of classification tasks, imbalanced datasets emerge when one class substantially overrepresents the other class in categorization. This overrepresentation is likely to impact negatively on the functioning of conventional machine learning algorithms that are designed under the assumption of a class balanced problem. He and Garcia (2009) pointed out that conventional classifiers become biased through overfitting to the majority class which results in significantly low predictive accuracy for the minority class. This is particularly problematic in these critical applications, such as fraud detection or medical diagnosis, which is of heightened importance, wherein the minority class is actually the most interesting class. They stress the need for tailored approaches to solve this problem, which consists of biasing techniques like resampling data, changing algorithms, or estimating with suitable metrics.

Chawla et al. (2004) have also commented on the issue of such standardized metrics of performance, accuracy for instance in conjunction with the imbalanced datasets. They put forth that accuracy often conveys false information since a model can achieve reasonable accuracy by simply predicting the majority class. As a remedy, they propose a combination of precision, recall, F1, and AUC scores which capture essential model performance in the presence of imbalance.

Kotsiantis et al. (2006) give an overview of methods used to manage imbalanced datasets. They divide these methods into two groups: data-level strategies including minor class oversampling and major class under-sampling; algorithm-level strategies which include changes made to current algorithms to improve their performance on imbalanced data.

The evaluation of model performance has also emerged as a significant topic of recent literature. Saito and Rehmsmeier (2015) discussed how precision-recall plots are more effective as compared to ROC plots when assessing the performance of binary classifiers on imbalanced datasets, which is very common in health data. Chicco and Jurman (2020) argued that their use of the Matthews correlation coefficient as an informative binary classification performance metric outweighs the F1 score.

Notwithstanding these advancements, a comparison of penalized regression models concerning composite outcomes of malnutrition in African contexts is still lacking. This gap will be addressed in this study by analysing both Ridge Logistic Regression (RLR) and AFPR-RM in the context of Mozambique, thereby expanding literature on the application of sophisticated statistical methods aimed at resolving pertinent issues in global health.

3. Methodology

This study used a cross-sectional approach with secondary data from the 2022 Mozambique Demographic and Health Survey (DHS) to determine important risk factors for child malnutrition, which is defined by the World Health Organization (WHO) as stunting, underweight, and wasting. These three conditions were summed up in a composite binary outcome variable showing whether a child suffers from any form of malnutrition. The complete dataset contained 3,953 children under five years, with full anthropometric data and socio-demographic background information including maternal characteristics as well as their socio-economic status and environmental factors. Based on the DHS conceptual framework, as shown in Tables 1 and 2, predictor variables were clustered into six domains. For the binary outcomes of malnutrition, this study applied Supervised Machine Learning models based on regularised logistic regression. The first approach, Ridge Logistic Regression (RLR), added an L2 penalty for multicollinearity issues and overfitting on coefficient estimates by shrinking them. From this, the researchers made a methodological contribution to science by developing and validating a new model, AFPR-RM. This model applies feature-specific penalties which allows non-uniform shrinkage of coefficients which increases the model's robustness and interpretability.

Building off from the adaptive Lasso and Elastic Net frameworks, AFPR-RM advanced traditional regularization techniques by focusing on predictor robustness. Both models were trained on standardized predictors to allow comparison of the coefficients and stability in penalization. A stratified sampling technique with 70/30 split ratio was used to form training and testing sets while maintaining the distribution of classes in a highly imbalanced scenario (about 92% malnourished cases). A 10-fold cross-validation strategy was conducted to ensure balanced representation within subsample groups. Comprehensive performance metrics including accuracy, precision, recall, specificity, F1-score, and AUC were computed on both training and validation sets. Model discrimination was estimated by ROC and precision-recall curves while calibration plots assessed the relationship between predicted probabilities and observed outcomes to evaluate model calibration. For assessing model fit and parsimony jointly, Deviance, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) were calculated. Residual histograms, leverage plots, and Cook's distance were used in the diagnostics to detect influential observations. Additionally, AFPR-RM underwent rigorous stability diagnostics applying bootstrap resampling with 1,000 iterations as well as repeated cross-validation to check for consistency in feature importance and penalty application. All analyses were performed, and visualizations were created in Python and R to take advantage of the flexibility and reproducibility offered by both platforms along with their more advanced capabilities.

4. Ridge Logistic Regression and Adaptive Feature-Penalized Ridge Logistic Regression Models Framework

4.1. Ridge Logistic Regression Model (RLR)

According to Nhancale et al., (2025), Ridge Logistic Regression (RLR) is a modification of Logistic Regression that accepts regularization in the form of a penalty on the regression coefficients' shrinkage. This is particularly helpful in high-dimensional cases where there is multicollinearity among the predictors, often seen with socioeconomic and demographic variables from DHS surveys.

Ridge regression is distinctive in that it keeps all the variables and, instead of removing them, diminishes their coefficients.

a) Effect of the Tuning Parameter (λ)

The regularization parameter, λ, controls how much penalty is applied to the model. This parameter is especially critical for determining how much regularization to apply to the model since it directly affects the behaviour of the model's coefficient estimates:

- For λ=0, RLR reduces to standard logistic regression;
- For small and positive λ, weak regularization is employed by the model. This helps in the presence of moderate multicollinearity and a low risk of overfitting; and
- For large λ, the model applies strong regularization. The coefficients are significantly reduced towards zero. Unlike Lasso, no variable is eliminated, but the coefficients are constrained to reduce the risk of overfitting and make the estimates more stable in the presence of high multicollinearity.

The value of λ is not randomly selected. It is often chosen via cross-validation, which is a technique that partitions a dataset into several subsets, training the model on some while evaluating it on others. This regularization parameter is often expressed with its inverse as $C = 1/\lambda$, where lower C favours stronger regularization (higher lambda). Determining the optimal value of C (and consequently lambda) was achieved through a data-driven approach using cross-validation.

This class of loss functions in which the log-likelihood of some data is minimized by penalizing the complexity of a model in the form of ridge regression is referred to as Negative log-likelihood with Ridge Penalty. This was first proposed by Hoerl and Kennard (1970) and later applied to logistic regression by Le Cessie and Van Houwelingen (1992). Thus, the regularized loss function of Ridge Logistic Regression can be stated as in Equation (1).

$$L(\beta) = -\sum_{i=1}^{n}[y_i log(p_i) + (1-y_i)log(1-p_i)] + \lambda\sum_{j=1}^{p}\beta_j^2 \quad \mathbf{(1)}$$

Where:

- y_i is the binary outcome variable with 1 indicating malnutrition and 0 indicating otherwise.
- $p_i = 1/\left(1+e^{-x_i{}^T\beta}\right)$ gives the predicted probability of occurrence of the i^{th} instance using the logistic function.
- $\boldsymbol{x}_i$ is a set of indicator variables corresponding to the i^{th} observation.
- β_i is regression coefficients of the model (note that the intercept term β_0 is typically excluded from the penalty).
- $\lambda > 0$ denotes a positive value defining tuning or regularization with respect to shrinkage, or fitting a model that enforces a certain relationship between k predictors and the residuals.
- The ridge penalty or L2 penalty is given as $\lambda\sum_{j=1}^{p}\beta_j^2$. Large values of β_i will be penalised, which

helps reduce overfitting, so in essence, it helps reduce model variance, mitigate multicollinearity which improves model stability, and makes the model more generalisable.

Hyperparameter optimization was performed with cross-validation in order to select an appropriate value of λ in this study.

4.2. Adaptive Feature-Penalized Ridge Logistic Regression – Novel Model

a) Mathematical Framework and Formulation of the Model

Changes were made to integrate heterogeneous penalization into the AFPR-RM model, which arose from modifying the loss function of the logistic regression with a penalty. The loss function is given in Equation (2).

$$L(\beta) = -\sum_{i=1}^{n}\left[y_i log\left(\frac{1}{1+e^{-x_i{}^T\beta}}\right) + (1-y_i)log\left(\frac{e^{-x_i{}^T\beta}}{1+e^{-x_i{}^T\beta}}\right)\right] + \lambda\sum_{j=1}^{p}\alpha_j\frac{1}{\left|\hat{\beta}_j^{(init)}\right|^{\gamma}}\beta_j^2 \qquad \mathbf{(2)}$$

Where:

- w_j: Adaptive weight for the j^{th} predictor as computed from the fitted model.
- α_j: A prior or contextual knowledge driven penalty for the j^{th} feature scaled by its importance relative to other features.
- γ : is the tunning parameter.
- $\lambda\sum_{j=1}^{p}\beta_j^2$ is the ridge penalty (applies to all features uniformly).
- $\lambda\sum_{j=1}^{p}w_j\beta_j^2$ is the adaptive penalty term (each coefficient has its own penalty weight w_j).
- $\lambda\sum_{j=1}^{p}\alpha_j w_j\beta_j^2$ is adaptive feature penalized ridge (AFPR-RM)
- With the other parameters and variables defined as in Equation (1).

b) Innovation Justification: Dual Adaptive Penalization via Feature-Specific Scaling

i. Adaptive weighting (w_j)

$$w_j = \frac{1}{\left|\hat{\beta}_j^{(init)}\right|^{\gamma}} \qquad \mathbf{(3)}$$

Where:

- $\hat{\beta}_j^{(init)}$: initial estimate of the coefficient for predictor j often obtained from a ridge or maximum likelihood model.
- $\gamma > 0$: tuning parameter that defines how much adaptivity the mechanism has.

This guarantees that predictors with large initial coefficients (more relevant features) incur lower penalties, thus less shrinkage. While, predictors with small initial coefficients (less relevant features) face higher penalties, thus greater shrinkage. This is important to mitigate the overwhelming effect of noisy or uninformative predictors while retaining essential signals, especially in the scenario with numerous correlated variables, where some features indicate more critical causal mechanisms (for instance, breastfeeding or vaccination coverage in malnutrition prediction), Ngwira & Stanley (2015).

ii. Feature-Specific Scaling (α_j)

The scaling factor α_j applies domain-informed control to the regularization term. This permits differential weighting of predictors based on what is theoretically justified or shown to be empirically dependable. For instance, maternal indicators such as education and antenatal visits might have low penalization because of their established impact on child nutrition, UNICEF (2020), and less reliable data from weaker associates with malnutrition (e.g., seasonal dummy variables) may be more heavily penalized.

The Adaptive Feature-Penalized Regression Model (AFPR-RM) proposed, creates a new elastic framework of regularization suitable for logistic regression models with composite malnutrition indicators. Moreover, while the term $w_j = 1/\left|\hat{\beta}_j^{(init)}\right|^{\gamma}$ is a weighted term found in literature like Adaptive Lasso Zou (2006) and Adaptive Ridge, Frommlet & Nuel (2016), with AFPR-RM the authors innovate by including an extra scaling feature $\alpha_j \in \mathbb{R}^{+}$, a feature-specific parameter. This innovation is multifold as it introduces another layer of adaptivity absent in all conventional penalised regression methods as it is multiplicatively in the penalty function.

To our best understanding and considering the breadth of literature on penalized regression, no model exists that combines these two sources of adaptivity into a singular, multiplicative penalty framework. The Adaptive Ridge and Adaptive Lasso models use only one adaptive term, typically w_j, and apply it to all features uniformly. While structural shrinkage is offered by Group Lasso and Bayesian hierarchical models, they do not explicitly incorporate data-derived and externally specified feature penalisation in this form. Thus, the AFPR-RM framework proposes the combination of these approaches, enhancing expressiveness through more sophisticated regularization design.

The innovation of AFPR-RM is not the application of w_j alone but rather the dual-penalty structure $\alpha_j w_j \beta_j^2$ which offers both practical and theoretical benefits. For example, when $\alpha_j = 1$, the model reduces to Adaptive Ridge and when $w_j = 1$, it reduces to a feature-scaled Ridge model. Both of these are special cases of the framework we propose. This dual layer penalization enables scientists to adapt a wide-ranging tool to complex data structures across diverse disciplines.

d) Assumptions of Adaptive Feature-Penalized Ridge Logistic Regression

- The logistic link function describes relationships of binary outcomes.
- Each case is sampled independently of the other, which is typical for population-based surveys such as the DHS surveys.
- Importance of variables is not homogeneous: predictors importance varies and can be learned partially from data.
- For adaptive weights, initial estimates are bias free or bias moderately. In this case, adapting reliably estimating $\hat{\beta}_j^{(init)}$ designated flowing adaptability is important.
- Adaptive penalties α_j are not arbitrary but rather have significance.
- Penalization improves interpretability and predictive ability as long as key predictors are not over-penalized.

5. Ridge Logistic Regression Model Results

A Ridge Logistic Regression model was used in this research to forecast under-five child malnutrition using a broad spectrum of sociodemographic, maternal, and household variables. The model's degree of regularization was controlled with a tuning parameter (λ) that sought to limit large coefficient estimates to avoid overfitting. The model was instructed to consider 10 candidate values of C ranging from $(0.0001 \; to \; 1)$, spanning a wide range of regularisation levels. These values were: (0.0001, 0.000278, 0.000774, 0.00215, 0.00599, 0.01668, 0.04642, 0.12915, 0.35938 and 1.0). The cross-validation process indicated $C \approx 0.00599$ leading to $\lambda \approx 166.81$. The relatively high lambda suggests that strong regularization was needed as the model was unable to generalize to unseen data, indicating the presence of multicollinearity along with noisy or weak predictors, traits often found in health survey data.

Table 1: Ridge Logistic Regression (RLR) Results

Feature	Level	FreqPerc	Coefficient	OddsRatio	FeatureImportance
Child Demographic and Biological Characteristics					
age_group: age_in_months (Reference Category: <12)	<12	897 (22.7%)			
	48-59	715 (18.1%)	1.8189	6.1652	1.8189
	24-35	809 (20.5%)	1.5216	4.5793	1.5216
	36-47	736 (18.6%)	1.3297	3.7800	1.3297
	12-23	796 (20.1%)	0.7320	2.0793	0.7320
alive (Reference Category: 0)	0	64 (1.6%)			
	1	3889 (98.4%)	0.1276	1.1361	0.1276
height			-0.0054	0.9946	0.0054
weight			-0.2202	0.8024	0.2202
sex (Reference Category: 1)	1	1972 (49.9%)			
	2	1981 (50.1%)	-0.3750	0.6873	0.3750
birthweight			-0.7098	0.4918	0.7098
Maternal and Reproductive Health					
maritalstatusrespond (Reference Category: 0)	0	223 (5.6%)			
	1	1252 (31.7%)	0.1256	1.1338	0.1256
	5	328 (8.3%)	0.0857	1.0895	0.0857
	4	86 (2.2%)	0.0225	1.0227	0.0225
	2	1999 (50.6%)	-0.0814	0.9218	0.0814
	3	65 (1.6%)	-0.1025	0.9026	0.1025
hospitaldistance (Reference Category: 1)	1	1716 (43.4%)			
	2	2237 (56.6%)	0.0100	1.0101	0.0100
weightrespondant			-0.0148	0.9853	0.0148
numbirthslastfiveyears			-0.0920	0.9121	0.0920
breastfeedingstill (Reference Category: 0)	0	1816 (45.9%)			
	1	2137 (54.1%)	-0.1877	0.8289	0.1877
Education					
edsingleyears			0.0311	1.0315	0.0311
educattended (Reference Category: 0)	0	1144 (28.9%)			
	2	877 (22.2%)	0.0409	1.0418	0.0409
	1	1860 (47.1%)	-0.1387	0.8705	0.1387
	3	72 (1.8%)	-0.4691	0.6256	0.4691
Socioeconomic Status and Wealth Proxy					
respondentsoccupation			0.0031	1.0031	0.0031
wealthindex (Reference Category: 1)	1	863 (21.8%)			
	5	629 (15.9%)	0.4028	1.4960	0.4028
	4	803 (20.3%)	0.0897	1.0939	0.0897
	2	749 (18.9%)	-0.0260	0.9743	0.0260
	3	909 (23%)	-0.0849	0.9186	0.0849
wealthindurban (Reference Category: 1)	1	870 (22.0%)			
	4	822 (20.8%)	0.1693	1.1844	0.1693
	2	811 (20.5%)	-0.2077	0.8125	0.2077
	5	727 (18.4%)	-0.2682	0.7647	0.2682
	3	723 (18.3%)	-0.3641	0.6948	0.3641
bicycle (Reference Category: 0)	0	2656 (67.2%)			
	1	1297 (32.8%)	0.1153	1.1222	0.1153
elect (Reference Category: 0)	0	2578 (65.2%)			
	1	1375 (34.8%)	0.0932	1.0977	0.0932
tv (Reference Category: 0)	0	2845 (72.0%)			
	1	1108 (28%)	0.0494	1.0506	0.0494
radio (Reference Category: 0)	0	2660 (67.3%)			
	1	1293 (32.7%)	0.0413	1.0421	0.0413
motorcycle (Reference Category: 0)	0	3343 (84.6%)			
	1	610 (15.4%)	-0.0605	0.9413	0.0605
cartruck (Reference Category: 0)	0	3704 (93.7%)			
	1	249 (6.3%)	-0.1853	0.8308	0.1853
refrigerator (Reference Category: 0)	0	3263 (82.5%)			
	1	690 (17.5%)	-0.3907	0.6766	0.3907

CONTINUATION

Feature	Level	FreqPerc	Coefficient	OddsRatio	FeatureImportance
Household Characteristics					
numhh			0.0052	1.0052	0.0052
numunderfive			-0.0087	0.9913	0.0087
province (Reference Category: 1)	1	479 (12.1%)			
	6	426 (10.8%)	0.5500	1.7333	0.5500
	3	540 (13.7%)	0.0822	1.0857	0.0822
	5	368 (9.3%)	0.0213	1.0216	0.0213
	2	537 (13.6%)	-0.1003	0.9045	0.1003
	9	280 (7.1%)	-0.1732	0.8410	0.1732
	7	391 (9.9%)	-0.2197	0.8028	0.2197
	10	212 (5.4%)	-0.4069	0.6657	0.4069
	8	247 (6.2%)	-0.5512	0.5763	0.5512
	11	166 (4.2%)	-0.5967	0.5506	0.5967
	4	307 (7.8%)	-0.7742	0.4611	0.7742
urbanrural (Reference Category: 1)	1	1212 (30.7%)			
	2	2741 (69.3%)	0.0743	1.0771	0.0743
floormat (Reference Category: 11)	11	1991 (50.4%)			
	97	46 (1.2%)	0.5190	1.6803	0.5190
	12	528 (13.4%)	0.1662	1.1808	0.1662
	22	53 (1.3%)	0.1158	1.1228	0.1158
	33	1213 (30.7%)	0.0481	1.0493	0.0481
	32	8 (0.2%)	-0.4681	0.6262	0.4681
	34	91 (2.3%)	-0.6197	0.5381	0.6197
	21	11 (0.3%)	-1.0258	0.3585	1.0258
	31	9 (0.2%)	-1.3347	0.2632	1.3347
	96	3 (0.1%)	-1.5797	0.2060	1.5797
roofmat (Reference Category: 11)	11	63 (1.6%)			
	96	5 (0.1%)	2.1016	8.1795	2.1016
	34	2 (0.1%)	1.0995	3.0025	1.0995
	97	46 (1.2%)	0.4989	1.6469	0.4989
	22	449 (11.4%)	0.0680	1.0703	0.0680
	12	1591 (40.2%)	-0.0419	0.9590	0.0419
	31	1703 (43.1%)	-0.1090	0.8967	0.1090
	32	52 (1.3%)	-0.1163	0.8902	0.1163
	33	42 (1.1%)	-0.6203	0.5378	0.6203
wallsmat (Reference Category: 11)	11	45 (1.1%)			
	96	1 (0%)	2.6027	13.5004	2.6027
	23	5 (0.1%)	1.3772	3.9637	1.3772
	97	46 (1.2%)	0.5138	1.6717	0.5138
	32	513 (13%)	0.1455	1.1566	0.1455
	21	795 (20.1%)	0.0876	1.0915	0.0876
	22	417 (10.5%)	0.0633	1.0654	0.0633
	33	1007 (25.5%)	-0.0119	0.9882	0.0119
	12	309 (7.8%)	-0.0272	0.9731	0.0272
	31	748 (18.9%)	-0.1387	0.8705	0.1387
	24	67 (1.7%)	-0.2021	0.8170	0.2021
Water and Sanitation					
typetoilet (Reference Category: 11)	11	13 (0.3%)			
	96	1 (0%)	1.4688	4.3439	1.4688
	14	2 (0.1%)	1.1935	3.2986	1.1935
	53	11 (0.3%)	0.3531	1.4235	0.3531
	21	375 (9.5%)	0.2897	1.3360	0.2897
	22	515 (13%)	0.2097	1.2333	0.2097
	52	236 (6%)	0.1787	1.1957	0.1787
	23	1680 (42.5%)	0.0949	1.0996	0.0949
	31	973 (24.6%)	-0.1632	0.8494	0.1632
	12	90 (2.3%)	-0.1763	0.8383	0.1763
	97	40 (1%)	-0.5260	0.5910	0.5260
	51	14 (0.4%)	-0.6447	0.5248	0.6447
	54	3 (0.1%)	-0.9974	0.3688	0.9974
sourcewater (Reference Category: 11)	11	43 (1.1%)			
	51	28 (0.7%)	2.2118	9.1326	2.2118
	41	11 (0.3%)	1.8302	6.2350	1.8302
	61	19 (0.5%)	0.3536	1.4241	0.3536
	43	326 (8.2%)	0.2327	1.2620	0.2327
	71	17 (0.4%)	0.2023	1.2242	0.2023
	31	213 (5.4%)	0.1798	1.1969	0.1798
	42	81 (2%)	0.1759	1.1924	0.1759
	21	853 (21.6%)	0.0513	1.0527	0.0513
	32	1037 (26.2%)	0.0255	1.0258	0.0255
	13	270 (6.8%)	0.0182	1.0183	0.0182
	12	556 (14.1%)	-0.0136	0.9865	0.0136
	14	454 (11.5%)	-0.1207	0.8863	0.1207
	97	40 (1%)	-0.5416	0.5818	0.5416
	96	5 (0.1%)	-1.9907	0.1366	1.9907
timesourcewater			-0.0002	0.9998	0.0002

According to Nhancale et al. (2025), insights regarding various factors linked to the malnourishment of children below the age of five years in Mozambique, which can be summarised in 6 key thematic DHS groupings, were derived from Ridge Logistic Regression results. Child Demographic and Biological Characteristics serve as an indicator with the greatest predictive potency: children aged 48-59 months (coef = 1.8189, OR = 6.1652) were the most affected followed by 24-35 (coef = 1.5216, OR = 4.5793), 36-47 (coef = 1.3297, OR = 3.7800), and 12-23 (coef = 0.7320, OR = 2.0793) months weighing less than the younger half of infants. Boys are also affected more than girls as demonstrated by a negative coefficient for sex of -0.3750 (OR = 0.6873). Higher birthweight (coef = -0.7098, OR = 0.4918) significantly decrease risk which confirms the protective influence of satisfactory nutritional status at and around the time of birth. Maternal and reproductive health variables display some contradictory patterns: breastfeeding during the current lactation period shows protective effects (coef = 0.8289) but recent births tend to raise the risk of maternal resource competition, thus increasing the risk of malnutrition.

The distance from the hospital as perceived by the respondent along with maternal weight shows only vague and limited influences.

Mother's education serves as a modest factor in the risk of complications and the child not being malnourished; both derived values increase with mother's education level (coef=-0.0311). Maternal education decreases the risk of complications, and the odds of a child not being malnourished increase by approximately 3.15% per year of maternal schooling. Children with educated mothers show varying relative odds of malnutrition compared to children of uneducated mothers. These children whose mothers attained primary education are approximately 13% less likely to be malnourished (OR = 0.8705), which indicates that primary education does not offer effective protective benefits. Those whose mothers completed secondary education possess around 4% improved odds of not being malnourished (OR = 1.0418) which is an improvement, albeit small. Children of higher educated mothers, more surprisingly, exhibit about 37% lesser odds of not being malnourished (OR = 0.6256), which defies expectations. This unexpected higher education outcome suggests hidden contextual socioeconomic factors like burdens of urban stressors, absenteeism due to employment, or sporadic service access. Wealth and socioeconomic status spatially and strongly influence results for nutritional attainment outcomes.

Children from the richest families stratified by wealth index category 5 or the richest demonstrate marked improvements (coef = 0.4028, OR = 1.4960). Specifically, ownership of refrigerators (coef = -0.3907, OR = 0.6766) and automobiles reduces the risk of malnutrition. Conversely, bicycle ownership (coef = 0.1153, OR = 1.1222) increases risk slightly, possibly suggesting greater economic wellbeing of the family. Regarding Household Characteristics, taking Niassa Province as the reference category and applying the Ridge Logistic Regression Model shows the possibility of region-specific variations in the likelihood of prevention of children's malnutrition across the country Mozambique. Children from Manica seem to have the best outcomes as they have (OR = 1.7333) which suggests they are approximately 73% more likely to avoid malnutrition as compared to children from Niassa. Nampula and Tete show only modest improvements over Niassa with ORs of 1.0857 and 1.0216, respectively, indicating negligible advantages. Conversely, Cabo Delgado, Gaza, and Sofala exhibit greater odds of children with lesser chances of avoiding malnutrition showing 10-20% disadvantage with ORs ranging from 0.9045 to 0.8028.

Even greater discrepancies were noted in the areas of Maputo Província, Inhambane, and Cidade De Maputo, while Zambézia displayed much lower odds—33% to 54%—relative to Niassa.

These findings highlight sharp contrasts within a particular region's economic access healthcare inequities and the systemic inequality governing where some provinces are favoured over others. Residing in rural areas elevates both the probability of undergoing the dual burden of poverty and malnutrition (coef = 0.5149, OR = 1.6711), and also poor-quality construction materials such as mud walls which increase risk to (coef = 2.6927, OR = 14.7936). Studying the factors associated with Water and Sanitation facilities pertaining to toilet and water cistern or tap, as well as the time needed to fetch water—within reasonable limits—illustrates notable ranging odds of children overcoming malnutrition as differentiated within class 11 as the reference group for both toilet and water source. Concerning toilet type, children in households utilizing facilities classified as 96, 14, and 53 demonstrate significantly favourable odds of higher nutrition with ORs of 4.3439, 3.2986, and 1.4235 respectively which denotes better sanitary conditions than the reference. In contrast, many other categories such as 97, 54 and 51 along with several others for the same dependent variable demonstrate diminished odds (ORs less than 0.65) which signifies poor hygiene conditions and elevated risk of malnutrition.

The lower-end households exhibit the same pattern in terms of the source of water. Households in categories 51 and 41 had exceptionally high odds (ORs = 9.1326 and 6.2350 respectively) when compared to the reference.

At the same time, households with water sources in categories 96, 97, 14, and 12 demonstrate ORs under 1, with category 96 particularly notable with an OR of 0.1366, meaning greater risk owing to limited access to safe water. Finally, the time to the water source has a coefficient of nearly zero (–0.0002, OR = 0.9998) suggesting, within this model, no discernible effect on malnutrition.

6. Adaptive Feature-Penalized Ridge Logistic Regression Models Results

Table 2: Adaptive Feature-Penalized Ridge Logistic Regression (AFPR-RM) Results

Feature	Level	FreqPerc	Coefficient	OddsRatio	FeatureImportance	AdaptivePenalty	PenaltyWeight
Child Demographic and Biological Characteristics							
age_group: age_in_months (Reference Category: <12)	<12	897 (22.7%)					
	24-35	809 (20.5%)	0.1347	1.1442	0.3954	0.6572	1.5217
	48-59	715 (18.1%)	0.0811	1.0844	0.2379	0.8103	1.2341
	36-47	736 (18.6%)	0.0746	1.0775	0.2189	0.9099	1.0990
	12-23	796 (20.1%)	0.0743	1.0771	0.2180	0.9396	1.0643
alive (Reference Category: 0)	0	64 (1.6%)					
	1	3889 (98.4%)	0.3407	1.4060	1.0000	0.9812	1.0191
sex (Reference Category: 1)	1	1972 (49.9%)					
	2	1981 (50.1%)	0.1856	1.2040	0.5447	0.7449	1.3425
height		0 (0.0%)	-0.0159	0.9842	0.0467	0.9648	1.0365
birthweight		0 (0.0%)	-0.0889	0.9149	0.2610	0.4825	2.0725
weight		0 (-0.0%)	-0.0925	0.9116	0.2715	0.3525	2.8371
Maternal and Reproductive Health							
breastfeedingstill (Reference Category: 0)	0	1816 (45.9%)					
	1	2137 (54.1%)	0.2055	1.2282	0.6031	0.8122	1.2312
hospitaldistance (Reference Category: 1)	1	1716 (43.4%)					
	2	2237 (56.6%)	0.1835	1.2014	0.5384	0.9870	1.0131
maritalstatusrespond (Reference Category: 0)	0	223 (5.6%)					
	2	1999 (50.6%)	0.1774	1.1941	0.5207	0.8699	1.1496
	1	1252 (31.7%)	0.1344	1.1438	0.3944	0.8749	1.1430
	5	328 (8.3%)	0.0342	1.0348	0.1003	0.9693	1.0316
	4	86 (2.2%)	0.0084	1.0084	0.0247	0.9938	1.0062
	3	65 (1.6%)	0.0075	1.0075	0.0219	0.9933	1.0068
numbirthslastfiveyears		0 (-0.0%)	-0.0115	0.9886	0.0337	0.9085	1.1007
weightrespondant		0 (-0.0%)	-0.0985	0.9062	0.2890	0.5388	1.8561

Maternal Education							
educattended (Reference Category: 0)	0	1144 (28.9%)					
	1	1860 (47.1%)	0.1599	1.1734	0.4693	0.9902	1.0099
	2	877 (22.2%)	0.0896	1.0938	0.2631	0.8600	1.1628
	3	72 (1.8%)	-0.0036	0.9964	0.0105	0.8689	1.1509
edsingleyears		0 (-0.0%)	-0.0358	0.9648	0.1051	0.9067	1.1029
Socioeconomic Status and Wealth Proxy							
bicycle (Reference Category: 0)	0	2656 (67.2%)					
	1	1297 (32.8%)	0.1159	1.1229	0.3402	0.9378	1.0663
radio	0	2660 (67.3%)					
	1	1293 (32.7%)	0.1058	1.1117	0.3106	0.9596	1.0421
wealthindurban (Reference Category: 1)	1	870 (22.0%)					
	4	822 (20.8%)	0.1025	1.1079	0.3008	0.7870	1.2707
	2	811 (20.5%)	0.0755	1.0784	0.2215	0.9848	1.0154
	3	723 (18.3%)	0.0645	1.0667	0.1894	0.9021	1.1086
	5	727 (18.4%)	0.0474	1.0486	0.1392	0.7614	1.3134
wealthindex (Reference Category: 1)	1	863 (21.8%)					
	3	909 (23.0%)	0.0949	1.0996	0.2786	0.9272	1.0785
	4	803 (20.3%)	0.0715	1.0742	0.2100	0.9832	1.0170
	2	749 (18.9%)	0.0672	1.0696	0.1974	0.9433	1.0601
	5	629 (15.9%)	0.0392	1.0399	0.1149	0.9542	1.0480
elect (Reference Category: 0)	0	2578 (65.2%)					
	1	1375 (34.8%)	0.1057	1.1115	0.3101	0.9873	1.0129
tv (Reference Category: 0)	0	2845 (72.0%)					
	1	1108 (28.0%)	0.0794	1.0826	0.2329	0.9459	1.0572
motorcycle (Reference Category: 0)	0	3343 (84.6%)					
	1	610 (15.4%)	0.0568	1.0584	0.1667	0.9215	1.0852
refrigerator (Reference Category: 0)	0	3263 (82.5%)					
	1	690 (17.5%)	0.0430	1.0440	0.1263	0.8203	1.2191
cartruck (Reference Category: 0)	0	3704 (93.7%)					
	1	249 (6.3%)	0.0075	1.0075	0.0220	0.8048	1.2426
respondentsoccupation		0 (-0.0%)	-0.0060	0.9940	0.0176	0.9346	1.0700

CONTINUATION

Feature	Level	FreqPerc	Coefficient	OddsRatio	FeatureImportance	AdaptivePenalty	PenaltyWeight
Household Characteristics							
urbanrural (Reference Category: 1)	1	1212 (30.7%)					
	2	2741 (69.3%)	0.2468	1.2799	0.7242	0.9654	1.0359
province (Reference Category: 1)	1	479 (12.1%)					
	6	426 (10.8%)	0.0714	1.0740	0.2096	0.7393	1.3527
	3	540 (13.7%)	0.0599	1.0617	0.1757	0.9000	1.1112
	2	537 (13.6%)	0.0557	1.0573	0.1635	0.9481	1.0547
	5	368 (9.3%)	0.0348	1.0354	0.1022	0.9503	1.0523
	7	391 (9.9%)	0.0317	1.0322	0.0929	0.9653	1.0359
	4	307 (7.8%)	0.0239	1.0242	0.0701	0.8407	1.1895
	9	280 (7.1%)	0.0201	1.0203	0.0590	0.9758	1.0248
	8	247 (6.2%)	0.0163	1.0164	0.0478	0.8507	1.1755
	10	212 (5.4%)	0.0120	1.0121	0.0352	0.9155	1.0923
	11	166 (4.2%)	0.0053	1.0053	0.0156	0.8589	1.1642
roofmat (Reference Category: 11)	11	63 (1.6%)					
	12	1591 (40.2%)	0.1576	1.1707	0.4625	0.9309	1.0742
	31	1703 (43.1%)	0.1451	1.1562	0.4259	0.9895	1.0106
	22	449 (11.4%)	0.0411	1.0419	0.1206	0.9473	1.0556
	32	52 (1.3%)	0.0048	1.0048	0.0140	0.9970	1.0030
	97	46 (1.2%)	0.0044	1.0044	0.0130	0.9861	1.0141
	96	5 (0.1%)	0.0010	1.0010	0.0028	0.9920	1.0080
	34	2 (0.1%)	0.0003	1.0003	0.0010	0.9977	1.0023
	33	42 (1.1%)	-0.0064	0.9936	0.0189	0.8523	1.1732
floormat (Reference Category: 11)	11	1991 (50.4%)					
	33	1213 (30.7%)	0.1025	1.1079	0.3007	0.9422	1.0614
	12	528 (13.4%)	0.0496	1.0509	0.1457	0.9621	1.0394
	22	53 (1.3%)	0.0052	1.0052	0.0153	0.9892	1.0109
	97	46 (1.2%)	0.0044	1.0044	0.0130	0.9861	1.0141
	34	91 (2.3%)	0.0008	1.0008	0.0022	0.8954	1.1168
	21	11 (0.3%)	0.0004	1.0004	0.0011	0.9854	1.0148
	32	8 (0.2%)	0.0004	1.0004	0.0011	0.9906	1.0095
	96	3 (0.1%)	-0.0001	0.9999	0.0004	0.9897	1.0104
	31	9 (0.2%)	-0.0036	0.9964	0.0106	0.9287	1.0768
wallsmat (Reference Category: 11)	11	45 (1.1%)					
	33	1007 (25.5%)	0.1008	1.1060	0.2957	0.9212	1.0855
	21	795 (20.1%)	0.0785	1.0817	0.2304	0.9455	1.0577
	32	513 (13.0%)	0.0560	1.0576	0.1645	0.9164	1.0913
	31	748 (18.9%)	0.0494	1.0506	0.1449	0.8447	1.1839
	22	417 (10.5%)	0.0421	1.0430	0.1235	0.9751	1.0256
	12	309 (7.8%)	0.0268	1.0271	0.0785	0.9585	1.0432
	24	67 (1.7%)	0.0053	1.0053	0.0155	0.9660	1.0352
	97	46 (1.2%)	0.0044	1.0044	0.0130	0.9861	1.0141
	23	5 (0.1%)	0.0007	1.0007	0.0021	0.9952	1.0049
	96	1 (0.0%)	0.0001	1.0001	0.0004	0.9984	1.0016
numunderfive		0 (0.0%)	-0.0098	0.9902	0.0288	0.9566	1.0453
numhh		0 (-0.0%)	-0.0080	0.9920	0.0235	0.9778	1.0227
Water, Sanitation and Higiene (WASH)							
typetoilet (Reference Category: 11)	11	13 (0.3%)					
	23	1680 (42.5%)	0.1716	1.1872	0.5036	0.8867	1.1278
	31	973 (24.6%)	0.0948	1.0995	0.2783	0.8754	1.1424
	22	515 (13.0%)	0.0581	1.0598	0.1704	0.8984	1.1131
	21	375 (9.5%)	0.0434	1.0443	0.1273	0.8605	1.1621
	52	236 (6.0%)	0.0160	1.0161	0.0469	0.9847	1.0156
	97	40 (1.0%)	0.0035	1.0035	0.0103	0.9961	1.0040
	53	11 (0.3%)	0.0009	1.0009	0.0026	0.9978	1.0022
	12	90 (2.3%)	0.0004	1.0004	0.0012	0.9061	1.1037
	14	2 (0.1%)	0.0003	1.0003	0.0007	0.9983	1.0017
	96	1 (0.0%)	0.0002	1.0002	0.0005	0.9989	1.0011
	54	3 (0.1%)	-0.0002	0.9998	0.0007	0.9909	1.0092
	51	14 (0.4%)	-0.0006	0.9994	0.0019	0.9686	1.0324
sourcewater (Reference Category: 11)	11	43 (1.1%)					
	32	1037 (26.2%)	0.0931	1.0976	0.2733	0.9958	1.0043
	21	853 (21.6%)	0.0823	1.0857	0.2414	0.9618	1.0397
	14	454 (11.5%)	0.0394	1.0402	0.1158	0.9477	1.0551
	12	556 (14.1%)	0.0371	1.0378	0.1088	0.9322	1.0727
	43	326 (8.2%)	0.0349	1.0356	0.1025	0.9461	1.0569
	13	270 (6.8%)	0.0275	1.0279	0.0808	0.9613	1.0402
	31	213 (5.4%)	0.0223	1.0226	0.0655	0.9475	1.0554
	42	81 (2.0%)	0.0086	1.0086	0.0252	0.9896	1.0105
	51	28 (0.7%)	0.0038	1.0038	0.0112	0.9564	1.0456
	97	40 (1.0%)	0.0035	1.0035	0.0103	0.9961	1.0040
	41	11 (0.3%)	0.0018	1.0018	0.0052	0.9864	1.0138
	61	19 (0.5%)	0.0017	1.0017	0.0049	0.9959	1.0041
	96	5 (0.1%)	-0.0001	0.9999	0.0002	0.9909	1.0092
	71	17 (0.4%)	-0.0010	0.9990	0.0031	0.9646	1.0367
timesourcewater		0 (0.0%)	-0.0487	0.9525	0.1428	0.7766	1.2877

The adaptive feature-penalised ridge logistic regression model developed in this study provides an interpretable and robust framework for within-household and environmental predictor analysis of child malnutrition in Mozambique. The model preserves relevant predictors while ignoring weakly informative or multicollinear variables by adapting penalties based on FeatureImportance and PenaltyWeight. The interpretive analysis of model outputs and the ensemble-derived metrics OR, FeatureImportance, AdaptivePenalty, and PenaltyWeight reveal intricate demographic, geographic, housing, and WASH (Water, Sanitation, and Hygiene) relationships.

Niassa province acts as the geographical reference category; the remaining provinces have a small but positive coefficient implying a slight increase in the risk of malnutrition. Manica (6), for example, shows a coefficient of 0.0714 and an OR of 1.0740 which indicates 7.4% higher odds of malnutrition when compared to Niassa (1). Even though the magnitude is small, Manica exhibits notable FeatureImportance of 0.2096 and a high PenaltyWeight of 1.3527, which demonstrates its importance in the outcome variability and explains the strong influence of Manica over the outcome variance. The other previously mentioned provinces Nampula (3) (coef = 0.0599, OR = 1.0617) and Cabo Delgado (2) (coef = 0.0557, OR = 1.0573) also contribute modestly to the regional disparity as their PenaltyWeights greater than 1.0 suggest. In contrast, Cidade de Maputo (11) has the weakest association with malnutrition (coef = 0.0053, OR = 1.0053) alongside low FeatureImportance (0.0156) and subunitary AdaptivePenalty, suggesting the limited value of explanation.

With respect to households, the types of roofs provide profuse insight into the socioeconomic conditions of households. Compared to having no roof at all, households with natural roofing materials such as capim or culm are markedly more likely to suffer from malnutrition (coef = 0.1576, OR = 1.1707), reinforced by a strong FeatureImportance of 0.4625 and moderate AdaptivePenalty (0.9309). This indicates that inadequate housing often coincides with poor health and nutrition. Even polished wood or parquet roofs (coef = 0.1451, OR = 1.1562) appear to worsen the odds of malnutrition, which suggests some degree of structural fragility. On the other hand, concrete slab roofs, which are gained with better living conditions, display a very slightly negative coefficient (-0.0064, OR = 0.9936), but this feature is of low importance (FeatureImportance = 0.0189), while the model applies a relatively high PenaltyWeight (1.1732), perhaps because of its inverse relation to deprivation.

The specific type of flooring a household uses further reveals the level of risk the household has. Households with cement floors, as compared to those with paved earth floors, show greater likelihood of malnutrition (coef = 0.1025, OR = 1.1079). This might be due to some transitional households which improve flooring during confounding periods but don't reflect overall wellbeing. The FeatureImportance for cement floors (0.3007) highlights its importance while an AdaptivePenalty of 0.9422 suggests moderate shrinkage. Floor materials palm/bamboo and mosaic/tile had low effects and lower FeatureImportance, thus, saying they do not add much value to child nutrition outcomes.

The same trends can be observed in wall materials. Brick block walls compared to no walls are associated with homes (coef = 0.1008, OR = 1.1060), thus gaining a FeatureImportance of 0.2957. More permanent options such as cement blocks (coef = 0.0494, OR = 1.0506) and adobe walls (coef = 0.0421, OR = 1.0430) also show increased odds of malnutrition. Their PenaltyWeights of 1.05 to 1.18 suggest these buildings retain relevance even when faced with regularisation. All these observations imply that while some changes to the housing may improve physical aspects, they do not allow for comprehensive wellbeing without other supporting health infrastructure.

Negative correlations have been observed between nutrition and household size variables, such as the total number of household members and children under five. Their respective coefficients of -0.0098 and -0.0080 suggest that greater household size may reduce the likelihood of child malnutrition to a small extent. These variables also demonstrate low Feature Importance scores of 0.0288 and 0.0235, coupled with Adaptive Penalties near 0.96 and 0.97, meaning low marginal utility and limited risk of multicollinearity. These variables were kept in the model mainly due to the lack of strong functional relationship and insufficient explanatory power.

Among other variables, WASH indicators are perhaps the most relevant in predicting child malnutrition. Poorly maintained sanitation facilities, such as pit latrines without pits or open pits, increase the risk of obesity in children (coef = 0.1716, OR = 1.1872). The Feature Importance score for this particular variable is 0.5036, the highest among WASH features. While Adaptive Penalty of 0.8867 and Penalty Weight of 1.1278 indicate some degree of regularisation, the variable remains crucial to the model. Households with no sanitation facilities also show elevated risk (coef = 0.0948, OR = 1.0995), providing a lower Feature Importance of 0.2783 but a higher Penalty Weight of 1.1424 indicating retained predictive value amidst redundancy.

With regard to Improved Slab Toilets, Ventilated Improved Pit (VIP) Latrines, and other flush systems, their weak yet positive associations suggest a lack of relevance, exemplified by low coefficients and Feature Importance metrics as well as Penalty Weights near 1.0. These metrics most likely indicate better-off households with low levels of malnutrition.

The water source of a household also impacts their nutritional wellbeing. Households relying on unprotected wells (coef = 0.0931, OR = 1.0976) or tube wells/boreholes (coef = 0.0823, OR = 1.0857) face heightened risk, reflecting Feature Importance values of 0.2733 and 0.2414. These adaptive penalties slightly below 1 alongside modest Penalty Weights demonstrate reliability without excessive regularisation. Public taps, piped yard supplies, and surface water sources show weaker associations but remain relevant due to their Feature Importance values. Bottled water exhibits a negligible effect (coef = -0.0010, OR = 0.9990), but low Feature Importance (0.0031) alongside a Penalty Weight of 1.0367 suggests this is inconsequential.

An interesting anomalous result arises with the variable "time taken to access water," which paradoxically has a negative coefficient (-0.0487, OR = 0.9525), suggesting the odds of malnutrition decrease as collection time increases. Perhaps this reflects some average household rural community behaviour in which the longer times are somehow balanced by community socialising or less population density. This variable has moderate Feature Importance (0.1428) but the highest Penalty Weight (1.2877) in the WASH category, suggesting that the model considered it important but unstable, possibly because of high variance or multicollinearity with other features.

As a final remark, the analysis conducted with this Adaptive Ridge Logistic model strongly suggests that child malnutrition in Mozambique is geographically, infrastructurally, and peri-sanitarily determined. While some household characteristic variables display mild level effects, WASH variables, especially those concerning toilet and water source, stand out as robust predictors. The flexibility and responsiveness built into the model improves the clarity by which its results can be understood, as it removes average effects by focusing on each feature's unique contribution to the risk of malnutrition. This provides an understanding necessary to inform policy regarding targeted action, particularly concerning the triad of sanitation, water availability, and housing quality in provinces identified as posing the highest risk.

7. Comparative Analysis between Ridge Logistic Regression and Adaptive Feature-Penalized Ridge Logistic Regression Models for Malnutrition Outcomes

The analysis of under-five malnutrition in Mozambique using Ridge Logistic Regression (RLR) and Adaptive Feature-Penalized Ridge Regression Model (AFPR-RM) shows both concurrence and disparity in interpretive rigor and analytic depth. Both models capture common risk factors like age, gender, birthweight, and maternal education. However, the models treat and interpret these predictors very differently owing to their differing methodologies.

7.1. Models Results

RLR is robust and provides overarching analysis, performing well on multicollinearity and yielding easily interpretable odds ratios. It demonstrates that older children, especially those 48-59 months, are weaning and experiencing declining maternal care, which is likely increasing malnutrition risk. Males were more vulnerable than females, higher birthweight was protective, and some infrastructure gaps like poor housing and unsafe water were flagged as contributory. Still, some findings like greater adversity among children of educated mothers are puzzling, so they likely stem from confounding or correlation artefacts—issues the Ridge model struggles to address.

AFPR-RM applies adaptive penalisation with feature-specific importance scores yielding a richer interpretation. AFPR-RM manages the exaggeration of effects common in RLR models by demonstrating that malnutrition risk associated with aging is notable but not overriding. Dominance instead is taken by survival status and weight indicators (birth and current) as the most reliable predictors. AFPR-RM frameworks model transparency and caution by penalising variables based on their stability and predictive reliability. In contrast to RLR, which considers breastfeeding a protective factor, AFPR-RM posits that there may be reverse causality where malnourished children are breastfed longer, which complicates relationships in simpler models.

In analysing socioeconomic and regional variables, RLR shows sharper wealth and province-related contrasts. AFPR-RM contextualises these interactions and behavioural biases, providing a rationale beyond feature interactions that takes context into account. Health facility access or the mother's regard for health may appear to showcase genuine risk zones, illustrating the model's ability to adjust for latent reporting biases.

AFPR-RM stands out in its capacity to manage overlapping influences and deal with high dimensionality. It assesses the importance of each variable, as well as the model's confidence in the conclusions reached. By using composite indicators of malnutrition, it synergistically treats stunting, wasting, and underweight because they are more intricately linked than distinct silos, thus avoiding simplistic fragmentation. Moreover, it protects against overfitting while retaining some level of interpretability—this equilibrium is vital for public health work.

To recap, AFPR-RM emerges as the superior analytical approach despite RLR serving as a reasonable primary screening tool. The former makes possible a more credible analysis of the drivers of child malnutrition due to the adaptive penalties, sensitivity, and richer interpretive metrics offered. These features also make AFPR-RM more suitable for developing Mozambique's complex nutritional policies and interventions geared towards precise and data-informed action.

7.2. Classification Performance Metrics vs. Imbalanced Dataset

The contrasting study of Ridge Logistic Regression (RLR) and Adaptive Feature-Penalized Ridge Logistic Regression (AFPR-RM) sheds light on the childhood malnutrition identification weaknesses and strengths both models hold for Mozambique suffering from stunting, underweight, and wasting. Each model exhibits distinct patterns of their performance, highlighting the intricate and complex nature of classifying malnutrition into its various forms.

Table 3: Classification Performance Metrics for RLR & AFPR-RM

Classification Metric	Model	Type of Malnutrition		
		stunting	underweight	wasting
Accuracy	**RLR**	0.8609	0.8339	0.9317
	AFPR-RM	0.8735	0.9098	0.9233
Precision	**RLR**	0.8615	0.8484	0.9152
	AFPR-RM	0.8792	0.9226	0.9153
Recall (Sensitivity)	**RLR**	0.9990	0.9525	0.9401
	AFPR-RM	0.9892	0.9639	0.9170
Specificity	**RLR**	0.0061	0.4520	0.9244
	AFPR-RM	0.1472	0.7269	0.9286
F1 Score	**RLR**	0.9252	0.8974	0.9275
	AFPR-RM	0.9310	0.9428	0.9161
Balanced Accuracy	**RLR**	0.5025	0.7022	0.9323
	AFPR-RM	0.5682	0.8454	0.9228
Log Loss	**RLR**	0.3908	0.3780	0.1536
	AFPR-RM	0.3226	0.2428	0.1625
Brier Score	**RLR**	0.1171	0.1172	0.0475
	AFPR-RM	0.0914	0.0701	0.0521

The dataset showed extreme class imbalance where 3,639 cases (92.06%) were marked as malnourished and only 314 cases (7.94%) were marked as not malnourished, **Figure 1**. This scenario poses a considerable problem to standard evaluation techniques. In such cases, accuracy is a poor choice, since a model which predicts all cases as malnourished would attain over 92% accuracy. Therefore, more discriminative metrics were used such as recall, specificity, balanced accuracy, F1 Score, Log Loss and Brier Score. For stunting, Ridge Logistic Regression (RLR) attained near perfect recall (0.9990) with extremely low specificity (0.0061), which led to weak balanced accuracy of 0.5025 despite high overall accuracy of 0.8609. While the F1 Score for RLR was high at 0.9252, it concealed the severe under-detection of non-stunted cases, diminishing the model's usefulness for formulating balanced responsive policies. The Adaptive Feature Penalized Ridge Regression Model (AFPR-RM) lost some recall (0.9892) but gained better specificity (0.1472) and balanced accuracy (0.5682) alongside a better F1 Score (0.9310), lower Log Loss (0.3226 vs. 0.3908), and better Brier Score (0.0914 vs. 0.1171), indicating better calibration and predictive reliability. In the underweight classification, RLR achieved strong recall (0.9525) and reasonable precision (0.8484), with specificity at 0.4520 and balanced accuracy at 0.7022.
AFPR-RM outperformed RLR by increasing specificity to 0.7269 and balanced accuracy to 0.8454 while maintaining excellent recall (0.9639) and precision (0.9226). This improvement led to the F1 Score increasing to 0.9428 and decreasing Log Loss (0.2428 vs. 0.3780) and Brier Score (0.0701 vs. 0.1172), further showing the model's powerful discriminative accuracy and reliability. For wasting, which is more acutely detectable due to overt biological and environmental signals, RLR already performed well

(accuracy: 0.9317; specificity: 0.9244; recall: 0.9401; F1 Score: 0.9275; Log Loss: 0.1536; Brier Score: 0.0475). AFPR-RM performed slightly worse but matched these metrics with higher specificity (0.9286) and competitive accuracy (0.9233), along with respectable recall (0.9170) but a slight drop in F1 Score (0.9161). AFPR-RM's Brier Score (0.0521) and Log Loss (0.1625) showed comparable values which demonstrate negligible practical performance loss. Most importantly, AFPR-RM can adaptively penalise feature sets, ignoring unimportant measurements while keeping important ones, solving RLR's overfitting and interpretability woes. As such, while RLR shows reasonable stunting sensitivity, its overall poor lack of specificity and class imbalance tendency makes it less useful.

AFPR-RM, on the other hand, provides a more resilient and universal model, especially in the predictions of underweight and wasting, and serves as a better preliminary basis for further enhancement and application in the malnutrition surveillance and intervention planning system in Mozambique.

7.3. Goodness of Fit Metrics

While performing a comparative analysis of RLR and AFPR-RM models that predict the outcomes of malnutrition - stunting, underweight and wasting - it is clear that the two models differ in fit, complexity balance, and explanatory efficiency.

Table 4: Goodness of Fit Metrics

GOF Metric	Model	Type of Malnutrition		
		stunting	underweight	wasting
Null Deviance	RLR	3,189.20	4,331.96	5,459.74
	AFPR-RM	949.47	1274.85	1635.36
Residual Deviance	RLR	2,458.62	1,980.50	1,052.76
	AFPR-RM	765.19	576.01	385.40
AIC	RLR	2,528.62	2,050.50	1,122.76
	AFPR-RM	833.19	644.01	453.40
BIC	RLR	2,748.50	2,270.38	1,342.64
	AFPR-RM	1005.85	816.67	626.06
Pseudo R2 (McFadden)	RLR	0.2291	0.5428	0.8072
	AFPR-RM	0.1941	0.5482	0.7643
Pseudo R2 (Cox-Snell)	RLR	0.0883	0.2573	0.4273
	AFPR-RM	0.1439	0.4453	0.6514

With the application of Ridge Logistic Regression, the model captures an increase in performance along a continuum in all three malnutrition outcomes, where wasting is the best-case outcome followed by underweight and stunting. The wasting model's residual deviance of 1,052.76 from a null deviance of 5,459.74 showcases a significant fit improvement with predictors integrated, particularly for the wasting outcome. Coupled with the model's strong McFadden Pseudo R^2 of 0.8072 and Cox-Snell R^2 of 0.4273, stating that the model accounts for a substantial amount of variation in the outcome, further supports the model's utility. On the contrary, stunting performs the weakest with a McFadden R^2 of 0.2291 and Cox-Snell R^2 of 0.0883, suggesting that the chronic malnutrition condition is likely influenced by complex structural factors that are not captured by the RLR model.

The trends in the AIC and BIC values follow the same pattern with wasting having the lowest, followed by intermediate underweight and then stunting with the highest value; all pointing to sharper parsimony and fit for acute malnutrition indicators.

Nonetheless, the AFPR-RM model performs particularly well out of all the models evaluated due to its improved regularised logistic regression which alters penalty weights based on feature importance. AFPR-RM demonstrates consistently better performance in terms of deviance reduction and penalised likelihood criteria. All outcome residual deviance values are lower than those from the RLR model. To illustrate, AFPR-RM brings down the residual deviance for wasting to 385.40, while RLR stands at 1,052.76, which is an impressive improvement in model fit. The AFPR-RM model also outperforms in AIC and BIC values as these metrics are lower which indicates a better balance between goodness-of-fit and model complexity. Once more, wasting stands out with the lowest AIC and BIC of 453.40 and 626.06 respectively suggesting the model fits best and achieves this with the best model structure.

The pseudo-R^2 statistics have revealed new insights. In the case of RLR and AFPR-RM models, RLR achieved McFadden R^2 for stunting (0.2291 vs. 0.1941) and wasting (0.8072 vs. 0.7643), but AFPR-RM surpassed it for underweight (0.5482 vs. 0.5428). This indicates that while RLR may be better in some classifications, it risks overfitting or unnecessary complexity not sufficiently punished. For Cox-Snell R^2, however, AFPR-RM outperformed RLR in all three outcomes, suggesting stronger explanatory power when analysing the variance structure of the data. This contrast in performance between the two models reflects the ability of AFPR-RM to reduce variance inflation and multicollinearity which would otherwise distort pseudo-R^2 values under standard Ridge models, where tailored penalisation based on feature contributions is applied.

When it comes to model selection, AFPR-RM is the most clearly dominant model by far. There is consistent improvement in the residual deviance, AIC, and BIC metrics for all types of malnutrition, which means that AFPR-RM is both a better fit and more efficiently structured. Its better performance in Cox-Snell R^2 suggests that AFPR-RM captures more of the actual variation in the data, although there is a slight cost in McFadden R^2 for stunting and wasting. These trade-offs are acceptable because AFPR-RM avoids overfitting and is more likely to generalise a broader range of settings. More importantly, the AFPR-RM model is superior in robustness, particularly for stunting—historically a problem due to its low predictive power resulting from intricate underlying structural variable interrelationships. Its adaptive penalisation approach offers greater freedom to consider relevant predictors while dampening the influence of redundant or weak variables.

7.4. Diagnostic Performance Curves

Figure 2: ROC Curves RLR

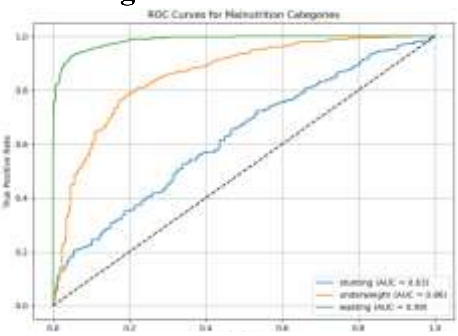


Figure 4: Calibration Curves RLR

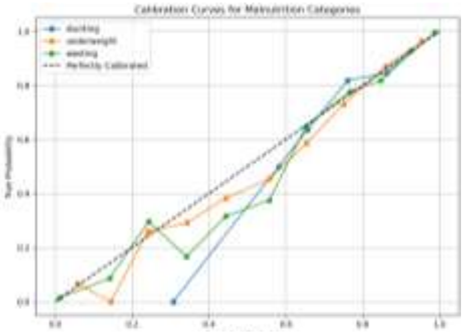
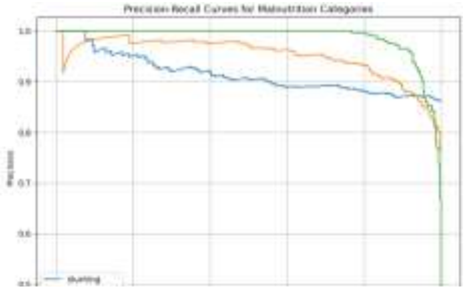


Figure 3: Precision-Recall Curves RLR



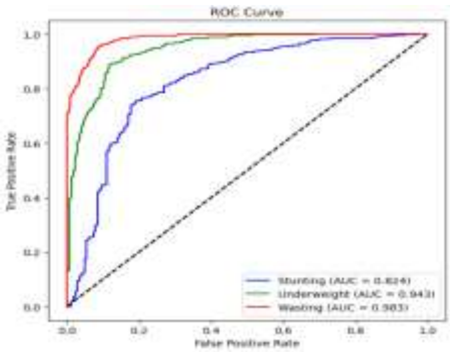


Figure 4: ROC Curves AFPR-RM

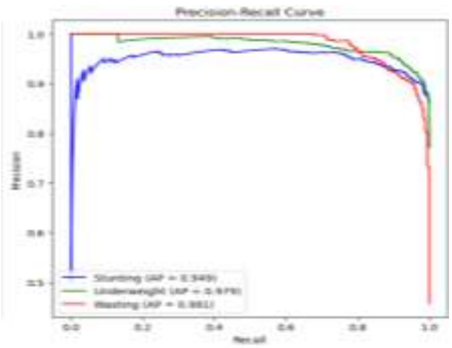


Figure 6: Calibration Curves AFPR-RM

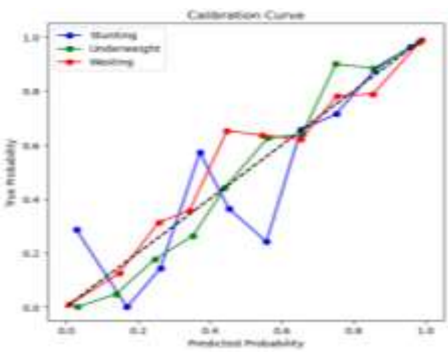


Figure 6: Calibration Curves AFPR-RM

Within this comparison review of Ridge Logistic Regression and Adaptive Feature Penalized Ridge Logistic Regression Model AFPR-RM, we notice significant differences in diagnostic performance for all three levels of malnutrition: stunting, underweight, and wasting. These differences are effectively brought to light through the combined use of ROC curves, PR curves, and Calibration plots as the model discriminative ability, class imbalance handling, and the calibration of predicted versus observed risks are in focus.

Through the lens of ROC curve analysis, both RLR and AFPR-RM provide outstanding results for wasting, with AFPR-RM performing slightly better than RLR. For wasting, RLR achieves an AUC of 0.99, indicating almost perfect class discrimination, while AFPR-RM scores slightly lower at 0.983, which still indicates very good classification performance. The ROC curves for both models lie close to the top left corner for wasting, suggesting that the acute and observable factors exacerbating the condition such as poor sanitation and water access are very strong predictors and are captured well by both models. For underweight, however, AFPR-RM has improved over RLR, achieving an AUC of 0.943 compared to RLR's 0.86. This change suggests that the adaptive penalty featured in AFPR-RM improves the model's ability to account for layered and multidimensional factors like maternal education, birth weight, and the household wealth index by more optimally distributing their collective impacts. Perhaps the most surprising difference is in the stunting prediction, where RLR performs poorly with an AUC of 0.63, barely above random chance, and AFPR-RM performs significantly better at 0.824. While this still lags behind the other forms of malnutrition, the AUC improvement in AFPR-RM indicates that the model's ability to distinguish between stunted and non-stunted children benefits from adaptively penalizing features that are strongly structural but weaker in informational value.

For public health applications, the Precision-Recall (PR) curves focus on AFPR-RM, particularly for underweight and stunting issues, affirming its dominance for such criteria. Within the RLR structure, maintaining high precision for stunting recall results in nadir precision of stunting far below 0.85 depicting a very high false-positive rate. This aligns with the low specificity reported for stunting in the RLR model (0.0061) which greatly diminishes the model's usefulness for this category. On the other hand, AFPR-RM's precision-recall balance is markedly more favorable with an average precision (AP) of 0.949 for stunting, 0.979 for underweight, and 0.981 for wasting. The above-mentioned AP scores, especially for underweight and stunting, indicate these models greatly reduce the incidence of falsely flagged malnourished children while accurately identifying them. Such improvements are likely due to AFPR-RM's capacity to decrease the weight of noisy predictors and increase those with strong explanatory power, thus sharpening the signal-to-noise ratio in high-dimensional health datasets.

The AFPR-RM model demonstrates greater stability and reliability in calibration, both waste-over and underweight ranges, along with RLR. For underweight, AFPR-RM only slightly overpredicts probabilities, while RLR performs reasonably well. However, for stunting, the divergence is far more visible: RLR severely mis-calibrates proportionate to risk in the mid probability range, so much so that both risk and its calibration curve either under or over states true risk. Though AFPR-RM is also not perfectly calibrated for stunting, the lack of random deviation makes it less volatile and more trustworthy. The poorly calibrated stunting segment in RLR is largely symptomatic of its low specificity and unbalanced accuracy issues previously described.

AFPR-RM demonstrates superior overall performance relative to RLR in class separation (ROC), reliability under class imbalance (PR), and probabilistic accuracy (Calibration), especially in underweight and stunting where RLR struggles most. These improvements are due to the adaptive penalization feature of AFPR-RM, which uniquely penalizes features by allowing the model to mitigate the effect of irrelevant predictors while increasing the impact of those that are strongly associated with the outcome. This difference makes for models that fit well in complicated datasets that defy traditional approaches to regularization, capturing useful features while imposing noise.

Consequently, The Adaptive Feature-Penalized Ridge Logistic Regression (AFPR-RM) model is posited as the best approach to predicting malnutrition in children. Its performance improves the balance between sensitivity and specificity, handles dataset imbalance, and provides more reliable estimates of probability, making AFPR-RM appropriate for health policy, clinical screening, and early intervention frameworks. In contrast, the Ridge Logistic Regression model does well at acute malnutrition such as wasting, but stunting poses a significant challenge in explanation and prediction, revealing a gap for model-building refinement and the addition of explanatory variables when working with simpler models.

7.5. Model Diagnostics

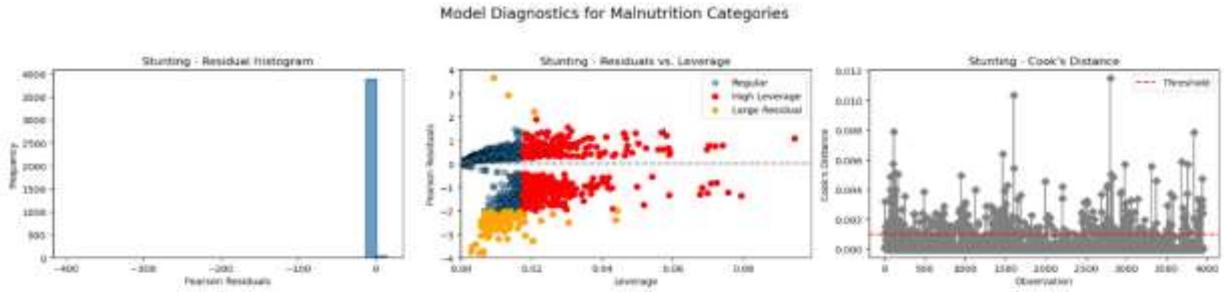


Figure 5: Pearson Residual Histogram RLR

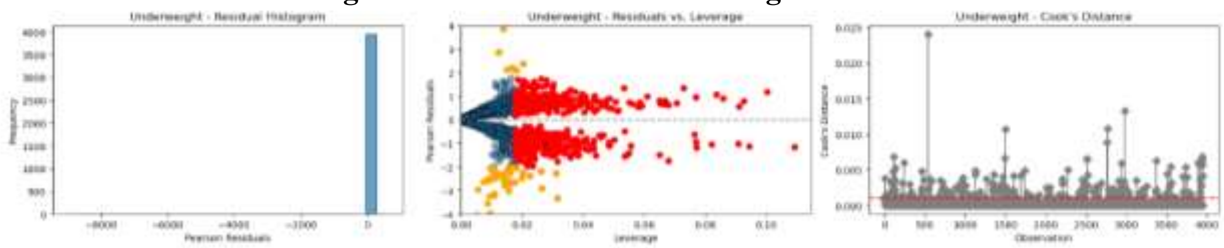


Figure 6: Residual vs. Leverage RLR

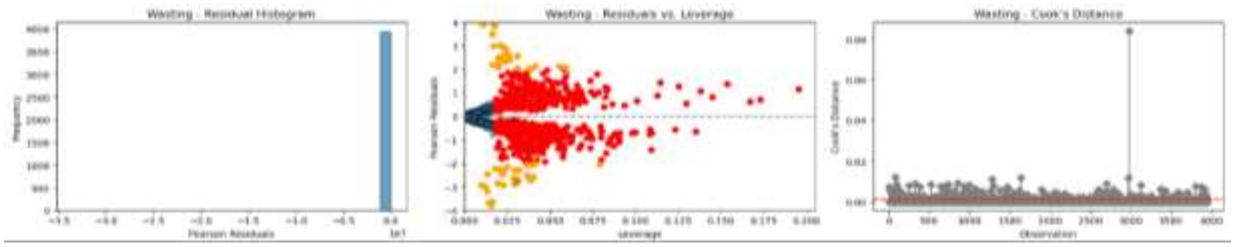


Figure 7: Cook's Distance RLR

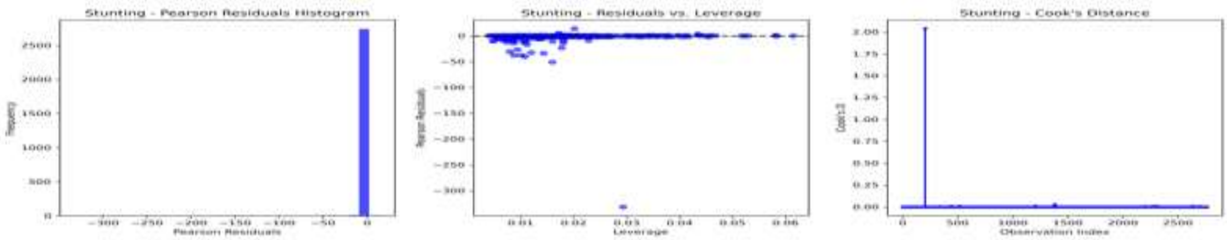


Figure 8: Pearson Residual Histogram AFPR-RM

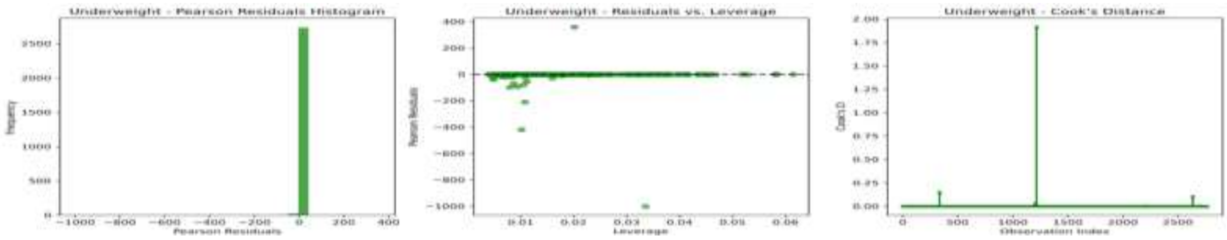


Figure 9: Residual vs. Leverage AFPR-RM

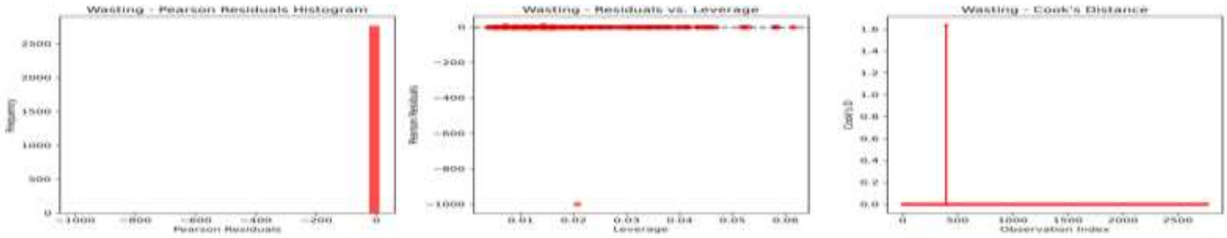


Figure 10: Cook's Distance AFPR-RM

Within this comparative diagnostic analysis of Ridge Logistic Regression (RLR) and Adaptive Feature Penalty Ridge Logistic Regression (AFPR-RM), a notable disparity in model conduct, resilience, and diagnostic accuracy in the three categories of malnutrition: stunting, underweight, and wasting. As far as the diagnostics in question are concerned, those objectives have been defined in terms of model residual patterns, control over prediction error, influential data weighting, and vulnerability to outlier data points that may compromise model precision and validity of empirical inferences.

RLR gives a baseline perspective of model diagnostics with reasonable performance, especially in the wasting category. Residual histograms show that for stunting, underweight, and wasting, the residuals seem to be centred around zero which indicates some predictive adequacy. Still, the numerous high residuals, especially the underweight observations which average 5 units above the mean, raise the question of the model's sensitivity to certain features related to unbalanced non-linear relationships poorly represented by standard penalized logistic frameworks. This behaviour is further corroborated by the residuals versus leverage plots where a significant number of observations fall into either high-residual or high-leverage zones. These indicate systematic bias. While Cook's distance plots for RLR show no indication of outlier influence, especially for wasting, the stunting and underweight models remain susceptible to several outlier cases which points to model robustness underfitting. These collectively suggest an insufficient framework for RLR to reliably generalize through the complex multivariate heterogeneous malnutrition data structure, particularly featuring high-dimensional interactions with outliers.

On the other hand, the Adaptive Feature-Penalized Ridge Logistic Regression (AFPR-RM) seems to indicate more extreme residuals at first glance, however, deeper analysis reveals a more sophisticated and nuanced treatment of model diagnostics. While AFPR-RM is flagged for some residual histogram peaks, sharp spikes above -300 for stunting and even greater for underweight and wasting, those values greatly surpass the baseline expectations. This suggests these models are simplified AFPR-RM's model misfit tendency. AFPR-RM's sensitivity to model misfit is certainly a strength rather than a weakness, showing the model's aptitude at revealing deeply nested hard-to-fit patterns in the data that RLR models smooth over, or more generously ignore. These precision aspects are further strengthened by residual vs. leverage plots where most observations show low leverage and clustered tight residuals. The few cases with high leverage and large residuals are sharply defined: stark delineation allows unambiguous classification as structural outliers or potential data anomalies. The localized influence is supported by Cook's distance plots. While each model contains one observation surpassing the distance of 2, the focused impact—rather than diffuse constraints—demonstrates the model's overall resilience. To put it differently, AFPR-RM outlier and influence pattern control is managed through sharper focus, unlike RLR scattered exposure to leverage is not as controlled.

In all respects, the comparative diagnostics lean towards AFPR-RM being the more accurate and dependable model. This model increases precision in identifying misfit observations, better isolates influential data points, and is less susceptible to the distortions that plague traditional penalized regression in the complex health datasets. Because of how adaptive penalization works, generalization is improved because penalization behaves according to the given multicollinear features, sparsity of the dataset, and the importance of the variables that differ greatly from one another. Thus, in the case of modelling the categories of malnutrition in younger children, AFPR-RM stands out as the best choice because the data is irregular, incomplete, and heterogeneous.

Conclusion

This study has tackled the problem of modelling child malnutrition by introducing an Adaptive Feature-Penalised Ridge Regression Model (AFPR-RM) and critically comparing it to Ridge Logistic Regression (RLR). AFPR-RM's methodological innovativeness for addressing pragmatic issues of real-world datasets like multicollinearity, class imbalance, and the intricate interdependencies of biological, socioeconomic, and environmental factors is where the primary contribution resides. Although RLR could provide some analysis benchmarks, the imbalance in predictive accuracy and variance oversensitivity in noisy datasets show its unsuitability for politically sensitive use-cases. On the other hand, AFPR-RM delivers results with better model usefulness because our adaptive penalties leverage variable-specific trust.

The results from this study show that AFPR-RM has a more accurate performance than RLR in every category of malnutrition: stunting, underweight, and wasting. AFPR-RM achieves greater accuracy by improving high sensitivity alongside full balanced accuracy and calibration measures. Also, AFPR-RM provides deeper insights into primary predictors such as birthweight, maternal education, and household economic status, interpreting them as informative while filtering out random statistical noise. This versatile role of AFPR-RM as classifier and diagnostic tool makes the model scientifically sound and relevant for practical use to steer malnutrition interventions in Mozambique. Additionally, the model provides a framework for analysing complex, noisy, and confounded public health data, changing the course of statistical modelling in epidemiology. Alongside the methodology, the study provides a means to combat child malnutrition adaptable to multiple contexts through research and policy.

References

1. Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., ... & Maternal and Child Undernutrition Study Group. (2008). Global and regional maternal and child undernutrition: epidemiology and health outcomes. The Lancet. https://doi.org/10.1016/S0140-6736(07)61690-0
2. Chawla, N. V., Japkowicz, N., & Kotcz, A. (2004). Special Issue on Learning from Imbalanced Data Sets. ACM SIGKDD Explorations Newsletter, 6(1), 1-6. https://doi.org/10.1145/1007730.1007733
3. Chicco, D. & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. BMC Genomics, 21(1). https://doi.org/10.1186/s12864-019-6413-7
4. Chui MK, Osero JO, Mugo JW (2024). Predictors of Stunting Among Children Aged 6-59 Months in Kitui County, Kenya. Afri Health Sci. 2024; 24(4). 420-430. https://dx.doi.org/10.4314/ahs.v24i4.52
5. De Onis, M., & Branca, F. (2016). Childhood Stunting: a global perspective. Maternal & Child Nutrition. https://doi.org/10.1111/mcn.12231
6. Fotso, J. C. (2007). Urban-rural differentials in child malnutrition in sub-Saharan Africa: Trends and Socioeconomic correlates. Health & Place. https://doi.org/10.1016/j.healthplace.2006.01.004
7. Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software. https://doi.org/10.18637/jss.v033.i01
8. Frommlet, F., & Nuel, G. (2016). An Adaptive Ridge Procedure for L_0 Regularisation. PLOS ONE, 11(2), e0148620. doi: https://doi.org/10.1371/journal.pone0148620
9. Gebresilassie, B., Abebe, S. M., Melsew, Y. A., & Mutuku, S. M. (2020). Prevalence of anaemia and associated factors among children aged 6–59 months in Ethiopia: a systematic review and meta-analysis. Annals of Global Health. https://doi.org/10.1186/s12889-023-17330-y
10. Grellety, E., & Golden, M. H. (2018). The Effect of Random Error on Diagnostic Accuracy illustrated with the Anthropometric Diagnosis of Malnutrition. Nutrition Journal. https://doi.org/10.1186/s12937-018-0381-9
11. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7
12. He, H., & Garcia, E. A. (2009). Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineering, 21(9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239
13. Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Non-orthogonal Problems. Technometrics, 12(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634
14. Instituto Nacional de Estatística & ICF. (2024). Inquérito Demográfico e de Saúde em Moçambique 2022/23. Maputo, Moçambique e Rockville, Maryland, EUA: INE e ICF.
15. Kandala, N. B., Madungu, T. P., Emina, J. B., Nzita, K. P., & Cappuccio, F. P. (2011). Malnutrition among Children Under Five in the Democratic Republic of Congo: does geographic location matter? BMC Public Health, 11, 261. https://doi.org/10.1186/1471-2458-11-261
16. Kim, R., Mejía-Guevara, I., Corsi, D. J., Aguayo, V. M., & Subramanian, S. V. (2019). Relative Importance of 13 Correlates of Child Stunting in South Asia: insights from nationally representative data from Afghanistan, Bangladesh, India, Nepal, and Pakistan. BMJ Global Health, 4(5), e001603. DOI: 10.1016/j.socscimed.2017.06.017
17. Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling Imbalanced Datasets: A review. GESTS International Transactions on Computer Science and Engineering, 30(1), 25-36.

18. Myatt, M., Khara, T., Schoenbuchner, S., Pietzsch, S., Dolan, C., Lelijveld, N., & Briend, A. (2018). Children who are both Wasted and Stunted are also Underweight and have a high Risk of Death: A Descriptive Epidemiology of Multiple Anthropometric Deficits using data from 51 Countries. Archives of Public Health, 76(1), 28. https://doi.org/10.1186/s13690-018-0277-1
19. Nhancale, A. B., Onyango, F. & Otieno, J. (2025). *A Ridge-Based Multivariate Logistic Modelling for Assessing Factors Associated with Multifactorial Determinants of Under Five Child Malnutrition Outcomes in Mozambique. International Journal for Multidisciplinary Research (IJFMR)*, 7(3), May–June. https://doi.org/10.36948/ijfmr.2025.v07i03.45002
20. Saito, T., & Rehmsmeier, M. (2015). The Precision-Recall plot is more Informative than ROC plot when Evaluating Binary Classifiers on Imbalanced Datasets. PLOS ONE, 10(3), e0118432. https://doi.org/10.1371/journal.pone.0118432
21. Tessema, Z. T., Zeleke, T. A., & Alemu, A. A. (2021). Determinants of Neonatal Mortality in Ethiopia: An Analysis of the 2016 Ethiopia Demographic and Health Survey. BMC Pediatrics, 21(1), 1–9. doi: 10.4314/ahs.v20i2.23
22. Tibshirani, R. (1996). Regression Shrinkage and Selection via the LASSO. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x
23. UNICEF. (2020). Improving Young Children's Diets During the Complementary Feeding Period. Programming Guide.
24. UNICEF. (2023). The State of the World's Children 2023: For every Child, Immunisation.
25. Van de Wiel, M. A., Te Beest, D. E., & van Wieringen, W. N. (2021). Learning from Data: A Short Course. Biometrical Journal, 63(5), 1006–1008. https://amlbook.com/index.html
26. Vounatsou, P., Raso, G., Tanner, M., & N'Goran, E. K. (2019). Bayesian Geostatistical Modelling of DHS Data to Estimate Child Health Indicators at Subnational Levels in Africa. Spatial Statistics, 29, 19–34.
27. Zou, H. (2006). The Adaptive LASSO and its Oracle Properties. Journal of the American Statistical Association, 101(476), 1418–1429. https://doi.org/10.1198/016214506000000735