International Journal on Data Science and Technology
2026, Vol. 12, No. 3, pp. 64–74
https://doi.org/10.11648/j.ijdst.20261203.12

Research Article

Machine Learning Approach to Assessment of Nutritional Status of Children with Cardiac Disease in Ethiopia

Tayu Nigusie Abebe*, Abdisa Jura Hunduma

Department of Statistics, Bule Hora University, Bule Hora, Ethiopia

Abstract

Nutritional status is an important determinant of health outcomes among children with cardiac disease, yet identifying children at risk of poor nutritional status remains challenging. Machine learning methods may provide useful tools for identifying important predictors and improving the classification of nutritional status in pediatric cardiac patients. The main objective of the study is to predict the nutritional status of children with cardiac disease using various machine learning methods, such as decision tree, logistic regression, random forest, support vector machine, and eXtreme Gradient Boosting. The predictive performance of each machine learning method was evaluated based on precision, accuracy, recall, f1 score, and the Kappa statistics value. The results indicate that the eXtreme Gradient Boosting method is recommended for predicting children nutritional status with precision, accuracy, recall, f1 score, and kappa statistics values: 0.7812, 0.7995, 0.7799, 0.7884, and 0.6299, respectively. The features that determine the nutritional status of children with cardiac disease are ROSS/NYHA, types of cardiac disease, pulmonary hypertension, anemia, pneumonia, and residence. This research will assist policy makers and healthcare providers to develop a framework for implementing necessary interventions and care practices to prevent complications associated with nutritional status in cardiac patients.

Keywords

Machine Learning, Nutritional Status, Cardiac, Children, Prediction

1. Introduction

Malnutrition is one of the common health problems experienced by children with cardiac disease in many countries [1, 2]. It can affect the overall growth and development of children [3]. Severe undernutrition affects 18.7 million, and moderate acute undernutrition affects 32.8 million children worldwide [2, 4]. Undernutrition is one of the most common causes of mortality and morbidity among children in the world [1, 3]. Assuming undernutrition as lack of available food due to poor economic status is not correct [5]. Cardiovascular disease is one of the common conditions which are associated with undernutrition [6, 7]. The prevalence of undernutrition among children with cardiac disease varies based on severity of the disease. According to [8] a study done in Nigeria, reported a prevalence of undernutrition (90.4%) and severe undernutrition (61.2%). However, there are limited works showing the prevalence of undernutrition among children with cardiac diseases from Ethiopia [6].

Cardiac disease and undernutrition is highly related in mul-

*Correspondence: Tayu Nigusie Abebe (tayutata2@gmail.com)

Received: 1 August 2026; **Accepted:** 4 September 2026; **Published:** 22 September 2026

tidirectional [8]. Undernutrition is associated with usual hospitalization, recurrent infections, and increased death rate [8, 9]. The cause of undernutrition in cardiac patients is due to different factors [9]. Cardiac diseases can worsen undernutrition due to different reasons; decreased food intake, increased energy requirement and venous congestion of the bowel [10, 11].

If not addressed in time, undernutrition in children can cause serious health challenges, increases the risk of developing long-term health challenges, such as high blood pressure, diabetes mellitus, behavioral disorders and death [12, 13]. It is necessary to detect the nutritional status of children cardiac patients quickly so that the stakeholder can provide the necessary treatment [14].

Different machine learning (ML) methods are used to predict nutritional status in children with cardiac disease, such as the k-nearest neighbor (kNN) [15, 16], naïve Bayes (NB) method [17], decision tree (DT) [18], logistic regression [19], support vector machine (SVM) [20], and learning vector quantization [20]. Different studies compare ML methods to predict malnutrition in patients [18, 19]. But, there is limited study in Ethiopia uses machine learning methods to predict nutritional statuses in children with cardiac disease. Therefore, this study aims to predict nutritional status of children with cardiac disease using different machine learning methods.

2. Methodology

2.1. Data Source

Nutritional status of cardiac children patient data were extracted from the cardiac center, Ethiopia. The study included cardiac patients of age below 18 years. Data of 323 patients comprise 141 boys and 182 girls were considered. The main purpose to collect data is to monitor and evaluate the nutritional status of cardiac children patients, including health status, causes of death, and more.

2.2. Study Variables

By reviewing the literature [21-24], list variables (Table 1) and extracted the selected variables from the cardiac patient information sheets.

Dependent variable: The study was focused on assessing the nutritional status of children with cardiac disease by utilizing the body mass index (BMI) as a measure. Nutritional status was assessed using height-for-age (HFA), weight-for-age (WFA), and weight-for-height (WFH), and interpreted in accordance with World Health Organization (WHO) [25]. According to WHO, undernutrition recommends a cut-off z-score of ≤ -2 to classify wasting, low HAZ (stunting) as moderate undernutrition, low WAZ (underweight) and and a z-score of ≤ -3 SD to define severe undernutrition [25]. However, for this study, overweight and healthy weight of children with cardiac disease was classified as a category of dependent variable.

Independent variable: Table 1 indicates the variable names, variable types, and categorizations based on previous literatures [21-24]. The independent variable includes; respondent's age, place of residence, types of cardiac disease, birth order, family history of cardiac disease, economic status, chamber enlargements, admission frequency, corrective surgery, pulmonary hypertension, anemia, New York Heart Association (ROSS/NYHA), heart failure, and gender were considered.

Table 1. *Variables and descriptions of the variables along with their categories.*

S/N	Variables	Variable type	Categories
1	Respondent's age (Month)	Categorical	< 24, 24-59, 60-143, >=144
2	Cardiac disease	Categorical	Congenital, Acquired
3	Residence	Categorical	Urban, Rural.
4	Birth order	Categorical	1^{st}, $2\text{-}3^{rd}$, >3
5	Family history of CD	Categorical	Yes, No
6	Economic status	Categorical	Below, average, above average
7	Chamber enlargements	Categorical	Yes, No
8	Pulmonary hypertension	Categorical	Yes, No
9	Corrective surgery	Categorical	Yes, No
10	ROSS/NYHA	Categorical	I, II, III, IV
11	Anemia	Categorical	Yes, No

S/N	Variables	Variable type	Categories
12	Pneumonia	Categorical	Yes, No
13	Heart failure	Categorical	No, Mild, Moderate to severe
14	Gender	Categorical	Boy, girl
15	Admission frequency	Categorical	No, 1, 2-3, >3

2.3. Data Processing

Many variables considered in this study were categorical, and a few of them were numeric. The numerical variable was converted into categorical variable for convenience of the study. Before model training, an exploratory analysis was conducted; categorical variables of the data were encoded for numerical values.

Before the data was randomly split into 20% data for testing and 80% data for training, it is important to clean the data from outliers, and duplicates (if any). The variables, gender, age, birth order, economic status, corrective surgery, and heart failure, will not be included in the classification process because they are not involved in the calculation. So, the remains variables from the data are used for testing and training [15]. The main variable is nutritional status, which is a binary variable that contains underweight and normal weight.

2.4. Feature Selection

We continue with variable selection after removing any missing values or duplicates. Variable selection was aims to reduce processing time and costs of computations [26]. To increase the predictive performance of the classification, we select variables that significantly contributed to the main class variable. We identify by conducting the chi-square test [27] between nutritional status and independent variables, and include those with a p-value < 0.05. Nine variables met these criteria and were used for developing the prediction model. These variables include ROSS/NYHA, types of cardiac disease, chamber enlargement, pulmonary hypertension, recent admission, admission frequency, anemia, pneumonia, and residence.

2.5. Dealing with Imbalanced Data

In this study on the nutritional data of children with cardiac disease, we noted a class imbalance, which result in biased estimates of measures of accuracy and precision. The percentage of health weight cardiac patient was 40.60%, which create an imbalanced distribution of the underlying classes and lead to biased and unreliable results while using machine learning.

2.6. Feature Transformation

Four feature transformation methods [27, 28] were applied (Table 2). From these methods, we select the best one for which the best machine learning can be extracted. The transformations we used are Min-Max Scaling (M1), log transformation (M2), square root transformation (M3), and Robust Scaling transformation (M4).

2.7. Machine Learning Algorithms

The research used five machine learning algorithms to predict the nutritional status (underweight, healthy weight) of cardiac patient in Ethiopia. The performance the algorithms was evaluated. The machine learning algorithms [19, 20, 29] used in this study include support vector machine (SVM), logistic regression (LR), random forest (RF), decision trees (DT), and extreme gradient boosting (XGB).

2.8. Performance Evaluation

Research used variety of measures to assess performance a models, as no single measure can fully capture all aspects of a model. Methods such as precision, accuracy, recall (sensitivity), f1 score, and the area under the curve were used to evaluate a model performance [18, 19].

2.9. Feature Importance

Identifying important variable or features is important to machine learning prediction. Feature importance rates illustrate the significance of each variable for decision-making. We have used Permutation Importance (PI), to identify the significant variables from the data [26, 27].

Ethics Approval and Consent to Participate

Ethical approval to conduct the studies was obtained from the cardiac center, Ethiopia with reference code (ref. no: chfe/9815/17/2024). Informed consent was obtained from guardian of patients under 18 years of age. The study strictly adheres to the principles outlined in the declaration of Helsinki.

2.10. Results

2.10.1. Descriptive Characteristics of the Study Participants

This study considered 323 children with cardiac disease (CD). Among those 182 (56.35%) were girls, and 56.04% were under age of five years at the study center. About 215 (66.56%) of the cardiac patients were resident in urban areas. Congenital cardiac disease (CCD) is more prevalent with 71.21% of patients diagnosed with CCD and 29.79% with acquired cardiac disease. Among the study patients, 35.29% had a history of anemia and 16.10% had history pulmonary hypertension. There are 47 (14.55%) corrective surgeries were performed, and 276 (85.45%) waiting to for corrective surgery during the study period. Moreover, 110 (34.05%) patients had no history of heart failure, 125 (38.69%) had mild cardiac failure and the others had moderate-to-severe cardiac failure.

Pneumonia was identified in 17.96% of cardiac patients, and anemia was also common with a prevalence of 35.29%. Looking into the family history of CD, 27 (8.36%) had a history of cardiac disease.

2.10.2. Machine Learning Algorithms Performance Evaluation

The study used four feature transformation (FT) techniques, square root, Min-Max Scaling, and Robust Scaling transformation (M1, M2, M3, and M4) along with five machine learning (ML) algorithms to classify the nutritional status of cardiac patients. The algorithms were evaluated based on various performance parameters, including kappa statistics, accuracy, precision, f1 score, recall, and AUC value. Tables 2-6 indicate algorithm's classification accuracy, kappa statistics, precision, f1 score, and recall values. Tables 7-10 also present the prediction results for underweight and healthy weight of cardiac patients, including precision, f1 score, AUC, and recall values.

Table 2. Accuracy (%) of machine learning algorithms for overall nutritional status of cardiac patients.

FT techniques	DTM	RFM	LRM	SVMM	XGBM
M1	65.424	75.807	56.127	56.454	78.925
M2	66.586	75.104	56.540	56.828	78.258
M3	67.213	75.929	57.605	57.444	79.106
M4	68.532	76.123	56.793	56.891	79.957

The results presents in Table 2 indicates that XGBM consistently perform over all other algorithms, achieving the highest accuracy across all feature transformation techniques. The highest value of predictive performance of 79.957% was using Robust Scaling (M4), suggesting that this transformation effectively handled outliers within the cardiac patient data.

Table 3. Kappa statistics (%) of machine learning algorithms for overall nutritional status of cardiac patients.

FT techniques	DTM	RFM	LRM	SVMM	XGBM
M1	47.345	59.945	31.154	31.932	62.543
M2	47.785	58.136	33.754	33.015	61.165
M3	45.654	60.778	32.462	34.551	62.264
M4	42.578	60.824	34.393	32.832	62.993

The results of analysis of Kappa statistics give in Table 3 indicate the reliability of the models beyond accuracy. XGBM have highest Kappa scores, with 62.993% when paired with Robust Scaling (M4). This indicates a significant level of agreement, validate that the model predictions are not the result of random chance.

Table 4. Precision (%) of machine learning algorithms for overall nutritional status of cardiac patient.

FT techniques	DTM	RFM	LRM	SVMM	XGBM
M1	65.243	73.182	52.523	51.834	75.821
M2	65.621	72.693	54.675	52.838	75.166
M3	64.023	74.549	53.332	54.653	76.987
M4	62.232	75.654	55.533	53.998	78.122

The result of precision analysis given in Table 4 indicates the superiority of the XGB algorithm in identifying nutritional status in cardiac patients. The algorithm achieved highest precision over other methods, having 78.122% with Robust Scaling (M4).

Table 5. Recall (%) of machine learning algorithms for overall nutritional status of cardiac patients.

FT techniques	DTM	RFM	LRM	SVMM	XGMM
M1	65.284	73.034	54.785	55.293	77.632
M2	65.623	72.355	56.332	55.822	76.712
M3	64.084	74.073	55.476	56.814	77.443
M4	62.187	75.277	56.833	55.823	77.996

The results presented in Table 5 identify all relevant cases of nutritional status among cardiac patients. Robust Scaling (M4) optimized other measures; the XGB algorithm achieved high recall value of 77.996% when paired with M1. This value minimizes false negatives and ensures that patient's nutritional interventions are correctly captured by the algorithms.

Table 6. F1 score (%) of machine learning algorithms for overall nutritional status cardiac patients.

FT techniques	DTM	RFM	LRM	SVMM	XGBM
M1	65.554	73.683	53.564	51.192	76.122
M2	65.733	73.038	55.380	51.774	75.465
M3	64.421	74.784	54.037	52.665	77.231
M4	62.582	76.017	56.142	52.783	78.485

The results in Table 6 shows the f1 score for the algorithms, and XGB algorithm paired with the M4 had the highest f1 score, at 78.485%.

Table 7. AUC (%) of machine learning algorithms for the underweight and healthy weight of cardiac patients.

Nutritional status	FT techniques	DTM	RFM	LRM	SVMM	XGBM
	M1	62.543	69.865	58.126	57.243	78.483
Underweight	M2	65.721	74.223	49.843	49.825	85.011
	M3	67.843	72.094	49.937	49.564	79.521

Nutritional status	FT techniques	DTM	RFM	LRM	SVMM	XGBM
	M4	61.845	69.987	54.063	53.562	89.389
Health weight	M1	67.842	83.743	64.663	63.602	89.382
	M2	69.963	79.511	55.128	55.128	90.463
	M3	66.785	77.384	55.427	55.253	90.552
	M4	69.998	83.742	60.432	59.362	89.385

The area under curve (AUC) given in Table 7 indicates the discriminative capacity of the algorithm for nutritional status, underweight and healthy weight.

Table 8. *Precision (%) of machine learning algorithms for the underweight and healthy weight of cardiac patients.*

Nutritional status	FT techniques	DTM	RFM	LRM	SVMM	XGBM
Underweight	M1	68.665	77.734	60.412	58.524	80.433
	M2	70.574	76.926	59.224	57.884	79.573
	M3	63.453	79.833	58.887	59.728	81.024
	M4	62.436	79.221	61.056	59.376	81.926
Healthy weight	M1	69.496	81.077	60.312	57.088	85.821
	M2	69.022	81.243	62.582	58.452	84.703
	M3	73.486	81.912	61.638	57.523	85.108
	M4	71.443	85.294	62.322	57.590	88.598

The XGB algorithm indicates high performance, particularly for the healthy weight, where it achieved its highest value AUC of 90.552% using the M3 transformation techniques (Table 7). This value shows an ability to distinguish healthy patients from other attributes. Figure 1 indicates the AUC results for different algorithms for both categories.

The precision analysis for the underweight and healthy weight classes, as summarized in Table 8. XGB outperformed other algorithm, having a high precision value of 81.926% for the underweight and 88.598% for the healthy weight. This suggests that the boosting significantly reduces of false-positive. Figure 2 shows the precision results for algo-rithms for both categories.

Table 9. *F1 score (%) of machine learning algorithms for the underweight and healthy weight of cardiac patients.*

Nutritional status	FT techniques	DTM	RFM	LRM	SVMM	XGBM
Underweight	M1	66.237	72.685	62.975	65.216	77.443
	M2	66.348	70.704	58.512	61.138	75.228
	M3	60.278	72.778	58.596	64.027	76.159
	M4	59.287	73.864	62.221	64.214	77.986
Healthy weight	M1	74.113	83.286	65.423	65.134	89.992
	M2	75.474	82.845	70.157	69.743	89.964
	M3	76.587	83.729	69.964	68.522	88.984

Nutritional status	FT techniques	DTM	RFM	LRM	SVMM	XGBM
	M4	73.926	86.296	67.984	66.283	88.623

The f1 score performance was given in Table 9, it was shown that the M4 transformation techniques under the XGB attained the highest f1 score of 77.986% for the underweight class and 89.992% for the healthy weight class for M1 transformation techniques. Figure 3 indicates that the f1 score for the algorithms for both classes.

Table 10. *Recall (%) of machine learning algorithms for the underweight and healthy weight of cardiac patients.*

Nutritional status	FT techniques	DTM	RFM	LRM	SVMM	XGBM
Underweight	M1	65.185	69.492	67.129	75.243	81.306
	M2	63.743	66.624	58.956	66.132	78.964
	M3	58.478	68.065	59.435	70.459	79.547
	M4	57.583	70.453	64.703	71.418	81.303
Healthy weight	M1	81.028	87.340	72.983	77.576	92.596
	M2	85.043	86.185	81.594	88.482	93.845
	M3	81.590	87.347	82.746	86.775	91.336
	M4	78.153	89.076	76.428	79.874	90.706

Lastly, Table 10 present the Recall (%) algorithm and M4 and M2 feature transformation methods yielded 81.303% and 93.845%, for underweight and healthy weight respectively. Figure 4 indicates the highest recall value for algorithms for both categories.

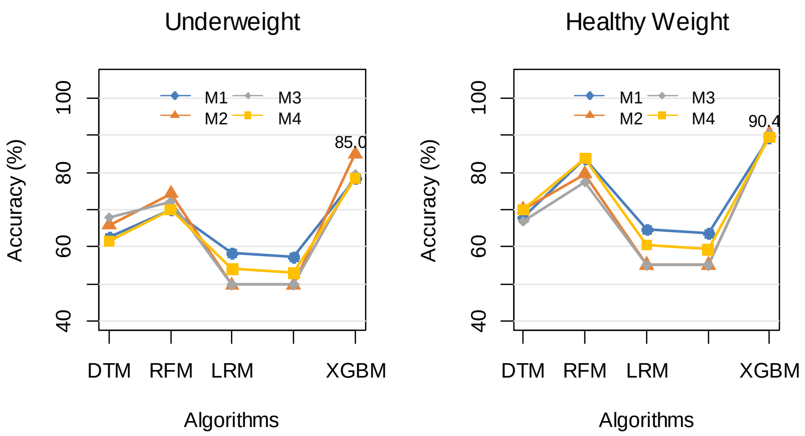

Figure 1. *AUC (%) of different algorithms for underweight and Healthy weight.*

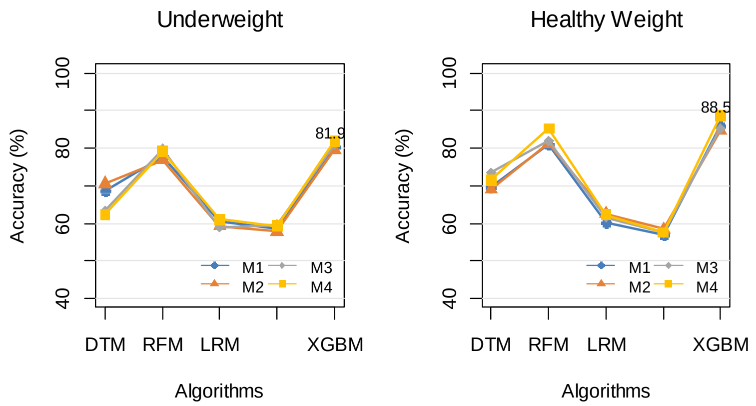

Figure 2. *Precision (%) of different algorithms for underweight and overweight/obese.*

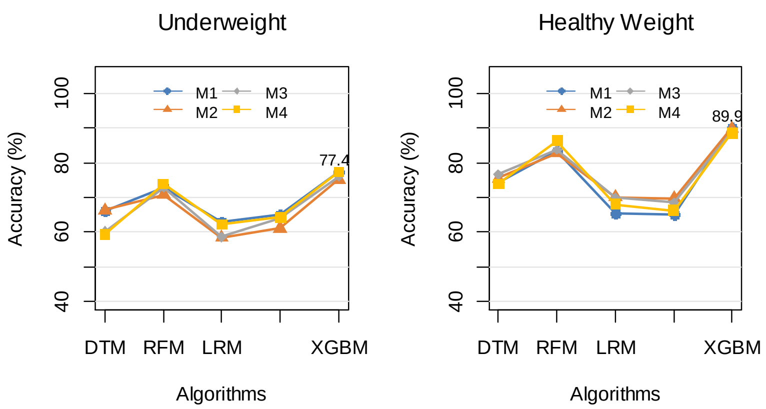

Figure 3. *F1 score (%) of different algorithms for underweight and healthy weight.*

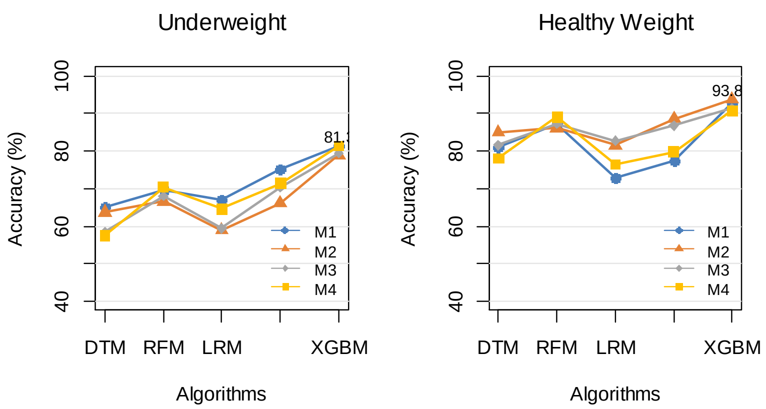

Figure 4. *Recall (%) of different algorithms for underweight and health weight.*

2.10.3. Variable Importance

After evaluating machine learning algorithm, feature importance approaches, PI was used for the XGB algorithm to rank the significant variables of the data. The variables, such as ROSS/NYHA, types of cardiac disease, chamber enlargement, pulmonary hypertension, recent admission, animia, pneumonia, residence, admission frequency were the most important variables of the nutritional status of cardiac patients (Figure 5 and Figure 6).

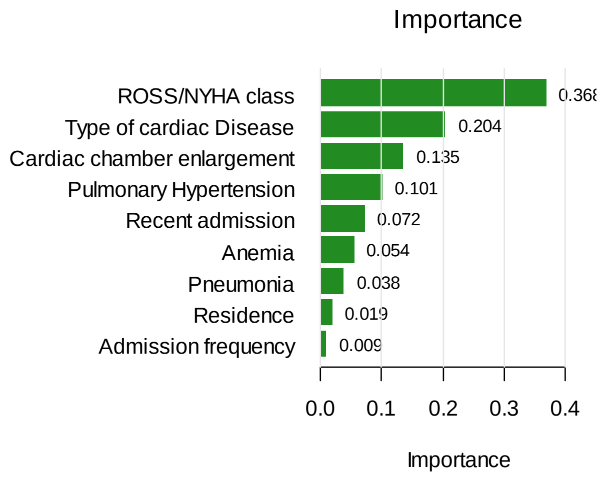

Figure 5. *Variable importance from the XGB model.*

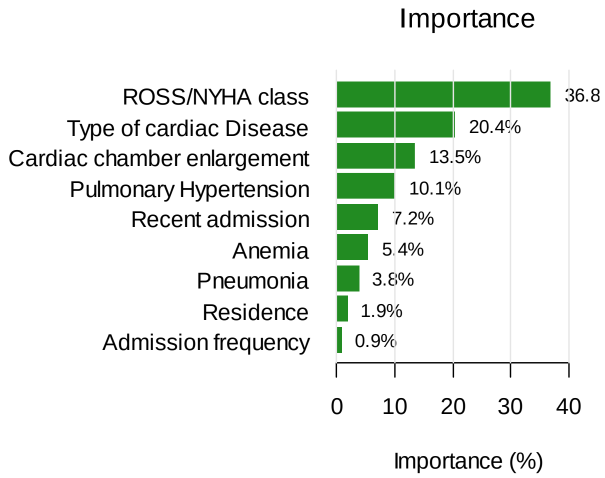

Figure 6. *Variable importance from the XGB model.*

3. Discussion

Many prediction models can identify the nutritional status of cardiac patients. But, there is lack of research on the use of machine learning approach to predict the nutritional status of cardiac patients in Ethiopia. The main objective of this study is to predict the nutritional status (underweight, healthy weight) of cardiac patient in Ethiopia. This study applied four different feature transformation techniques and five machine learning algorithms such as decision tree, logistic regression, random forest, support vector machine, eXtreme Gradient Boosting [15-17].

The study by [30], the impact of various feature transformation techniques on cardiac failure data was examined and used logistic regression, eXtreme Gradient Boosting, decision tree, and random forest model with scaling methods; standard, Max-Abs, Min-Max, quantile, and robust transformation [31]. The study indicates that XGB better performance with standard and robust transformation, which is consistent with our findings.

The XGB algorithm has the best prediction accuracy and the highest AUC score compared to other machine learning models for nutritional status pregnant women's [32]. The study by [33] found that the ADB, RF, and XGB algorithms were an effective at predicting nutritional status of women at childbearing age's [34], supporting this study's findings.

The study by [35] used XGB and Random Forest (RF) to analyze risk factors for obesity and overweight women. They found that the XGB algorithm is the best with highest accuracy and f1 score, and align with the study.

In addition to identifying the best predictive models, this study also determined the essential features predicting nutritional status among cardiac patient based on the best algorithm found in the study. Based on the important feature score for XGB algorithms suggested that the ROSS/NYHA, types of cardiac disease, pulmonary hypertension, recent admission, anemia, pneumonia, residence and admission frequency are the most important features for predicting the nutritional status of cardiac patients in Ethiopia. ROSS/NYHA is a significant factor in determining nutritional status of cardiac patients. This research aligns with previous studies [36]. The study revealed that anemia is a risk factor for developing malnutrition.

4. Rationale and Limitations of the Study

This study has limitations as it relies on single center data, which restricts its generalization. The analysis did not adjust for the sampling weight. Despite these limitations, the study's strength lies in identifying the best machine learning models using various performance evaluation methods, which is a main contribution to the current research.

5. Conclusions

This objective of this study was to assess the effectiveness of different machine learning models in predicting the nutritional status of children with cardiac disease in Ethiopia. We applied feature transformation techniques to the data and use different algorithms to analyze the data and investigate their performance. The XGB algorithms have best performance found in this study. The XGB algorithm, an important features that determine the nutritional status of children with cardiac disease, and ROSS/NYHA, types of cardiac disease, pulmonary hypertension, recent admission, anemia, pneumonia, residence and admission frequency are determinant factors for nutritional status. This research will assist policy-makers and healthcare providers to develop a framework for implementing necessary interventions and care practices to prevent complications and reduce the burden of nutritional status associated with cardiac disease.

Abbreviations

NYHA	New York Heart Association
SVM	Support Vector Machine
LR	Logistic Regression
RF	Random Forest
DT	Decision Trees
XGB	Extreme Gradient Boosting
PI	Permutation Importance
FT	Feature Transformation
ML	Machine Learning
AUC	Area Under Curve

Author Contributions

Tayu Nigusie Abebe: Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Software, Writing – original draft, Writing – review & editing

Abdisa Jura Hunduma: Conceptualization, Writing – review & editing

Data Availability Statement

The data supporting the outcome of this research work has been reported in this manuscript.

Conflicts of Interest

The authors declare no conflicts of interests.

References

[1] S. A. Khan, M. Fayaz, S. Hussain, S. A. Naqvi, O. bin Nazeer, and M. B. Khawaj, "Impact of Malnutrition: Studying the Effects of Undernutrition and Micronutrient Deficiencies on Cardiac Health in Children," *Indus J. Biosci. Res.*, vol. 3, no. 1, pp. 581-588, 2025. https://doi.org/10.70749/ijbr.v3i1.552

[2] A. Tabib, M. Aryafar, and B. Ghadrdoost, "Prevalence of malnutrition in children with congenital heart disease," 2019. https://doi.org/10.5812/compreped.84274

[3] J. Diao *et al.*, "Prevalence of malnutrition in children with congenital heart disease: a systematic review and meta-analysis," *J. Pediatr.*, vol. 242, pp. 39-47, 2022. https://doi.org/10.1016/j.jpeds.2021.10.065

[4] Dipasquale, V., Cucinotta, U., & Romano, C. (2020). Acute malnutrition in children: pathophysiology, clinical effects and treatment. *Nutrients*, *12*(8), 2413. https://doi.org/10.3390/nu12082413

[5] I. K. Murni *et al.*, "Outcome and factors associated with undernutrition among children with congenital heart disease," *PLoS One*, vol. 18, no. 2, p. e0281753, 2023. https://doi.org/10.1371/journal.pone.0281753

[6] M. Asrade, A. Shehibo, and Z. Tigabu, "Magnitude of undernutrition and associated factors among children with cardiac disease at University of Gondar hospital, Ethiopia," *BMC Nutr.*, vol. 7, no. 1, p. 43, 2021. https://doi.org/10.1186/s40795-021-00448-4

[7] H. Kwon *et al.*, "Incidence of cardiovascular disease and mortality in underweight individuals," *J. Cachexia. Sarcopenia Muscle*, vol. 12, no. 2, pp. 331-338, 2021. https://doi.org/10.1002/jcsm.12682

[8] X. Ruan *et al.*, "Associated factors of undernutrition in children with congenital heart disease: a cross-sectional study," *Front. Pediatr.*, vol. 12, p. 1167460, 2024. https://doi.org/10.3389/fped.2024.1167460

[9] M. Charkiewicz, Z. B. Wojszel, A. Kasiukiewicz, L. Magnuszewski, and A. Wojszel, "Association of Chronic Heart Failure with Frailty, Malnutrition, and Sarcopenia Parameters in Older Patients—A Cross-Sectional Study in a Geriatric Ward," *J. Clin. Med.*, vol. 12, no. 6, p. 2305, 2023. https://doi.org/10.3390/jcm12062305

[10] Yirdaw, B. W., Agmas, M. A., Mekonen, F. A., & Worku, S. Y. (2026). Nutritional status and associated factors in Ethiopian children with congenital heart disease: a systematic review and meta–analysis. *BMC pediatrics*, *26*(1), 315. https://doi.org/10.1186/s12887-026-06707-8

[11] J. D. Otvos *et al.*, "Multimarkers of metabolic malnutrition and inflammation and their association with mortality risk in cardiac catheterisation patients: a prospective, longitudinal, observational, cohort study," *lancet Heal. Longev.*, vol. 4, no. 2, pp. e72-e82, 2023. https://doi.org/10.1016/S2666-7568(22)00288-7

[12] A. Laillou, K. Baye, Z. Meseret, H. Darsene, A. Rashid, and S. Chitekwe, "Wasted children and wasted time: a challenge to meeting the nutrition sustainable development goals with a high economic impact to Ethiopia," *Nutrients*, vol. 12, no. 12, p. 3698, 2020. https://doi.org/10.3390/nu12123698

[13] J. C. Wells *et al.*, "The double burden of malnutrition: aetiological pathways and consequences for health," *Lancet*, vol. 395, no. 10217, pp. 75-88, 2020. https://doi.org/10.1016/S0140-6736(19)32472-9

[14] J. C. K. Wells *et al.*, "Beyond wasted and stunted—a major shift to fight child undernutrition," *Lancet Child Adolesc. Heal.*, vol. 3, no. 11, pp. 831-834, 2019. https://doi.org/10.1016/S2352-4642(19)30244-X

[15] G. Kocher and G. Kumar, "Machine learning and deep learning methods for intrusion detection systems: recent developments and challenges," *Soft Comput.*, vol. 25, no. 15, pp. 9731-9763, 2021. https://doi.org/10.1007/s00500-021-05893-0

[16] J. Padarian, B. Minasny, and A. B. McBratney, "Machine learning and soil sciences: A review aided by machine learning tools," *Soil*, vol. 6, no. 1, pp. 35-52, 2020. https://doi.org/10.5194/soil-6-35-2020

[17] A. F. A. H. Alnuaimi and T. H. K. Albaldawi, "An overview of machine learning classification techniques," in *BIO Web of Conferences*, EDP Sciences, 2024, p. 133. https://doi.org/10.1051/bioconf/20248401033

[18] O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, S. R. Meena, D. Tiede, and J. Aryal, "Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection," *Remote Sens.*, vol. 11, no. 2, p. 196, 2019. https://doi.org/10.3390/rs11020196

[19] A. M. Rahmani *et al.*, "Machine learning (ML) in medicine: review, applications, and challenges," *Mathematics*, vol. 9, no. 22, p. 2970, 2021. https://doi.org/10.3390/math9222970

[20] J. Wiens and E. S. Shenoy, "Machine learning for healthcare: On the verge of a major shift in healthcare epidemiology," *Clin. Infect. Dis.*, vol. 66, no. 1, pp. 149–153, 2018. https://doi.org/10.1093/cid/cix731

[21] T. Tsega, T. Tesfaye, A. Dessie, and T. Teshome, "Nutritional assessment and associated factors in children with congenital heart disease—Ethiopia," *PLoS One*, vol. 17, no. 9, p. e0269518, 2022. https://doi.org/10.1371/journal.pone.0269518

[22] R. Woldesenbet, R. Murugan, F. Mulugeta, and T. Moges, "Nutritional status and associated factors among children with congenital heart disease in selected governmental hospitals and cardiac center, Addis Ababa Ethiopia," *BMC Pediatr.*, vol. 21, no. 1, p. 538, 2021. https://doi.org/10.1186/s12887-021-02996-0

[23] S. R. Ismail, A. Mehmood, N. Rabiah, R. M. Abu-sulaiman, and M. S. Kabbani, "Impact of the nutritional status of children with congenital heart diseases on the early post-operative outcome," *Egypt. Pediatr. Assoc. Gaz.*, vol. 69, no. 1, p. 28, 2021. https://doi.org/10.1186/s43054-021-00076-2

[24] K. Tesfaye and T. Tsega, "Short-term impact of cardiac intervention on the nutritional status of malnourished children with congenital heart disease-a report from a developing African country, Ethiopia," *Pediatr. Heal. Med. Ther.*, pp. 465-475, 2023. https://doi.org/10.2147/PHMT.S433966

[25] M. Kamruzzaman *et al.*, "The anthropometric assessment of body composition and nutritional status in children aged 2-15 years: A cross-sectional study from three districts in Bangladesh," *PLoS One*, vol. 16, no. 9, p. e0257055, 2021. https://doi.org/10.1371/journal.pone.0257055

[26] D. Theng and K. K. Bhoyar, "Feature selection techniques for machine learning: a survey of more than two decades of research," *Knowl. Inf. Syst.*, vol. 66, no. 3, pp. 1575-1637, 2024. https://doi.org/10.1007/s10115-023-02008-7

[27] S. Salcedo-Sanz, L. Cornejo-Bueno, L. Prieto, D. Paredes, and R. García-Herrera, "Feature selection in machine learning prediction systems for renewable energy applications," *Renew. Sustain. Energy Rev.*, vol. 90, pp. 728-741, 2018. https://doi.org/10.1016/j.rser.2018.04.008

[28] A. Kaur, K. Guleria, and N. K. Trivedi, "Feature selection in machine learning: Methods and comparison," in *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, IEEE, 2021, pp. 789-795. https://doi.org/10.1109/ICACITE51222.2021.9404707

[29] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70-79, 2018. https://doi.org/10.1016/j.neucom.2017.11.077

[30] Y. Wang *et al.*, "Clinical prediction of heart failure in hemodialysis patients: based on the extreme gradient boosting method," *Front. Genet.*, vol. 13, p. 889378, 2022. https://doi.org/10.3389/fgene.2022.889378

[31] X. Xie *et al.*, "Early prediction of left ventricular reverse remodeling in first-diagnosed idiopathic dilated cardiomyopathy: a comparison of linear model, random forest, and extreme gradient boosting," *Front. Cardiovasc. Med.*, vol. 8, p. 684004, 2021. https://doi.org/10.3389/fcvm.2021.684004

[32] L. Liu *et al.*, "Early prediction of clinical scores for left ventricular reverse remodeling using extreme gradient random forest, boosting, and logistic regression algorithm representations," *Front. Cardiovasc. Med.*, vol. 9, p. 864312, 2022. https://doi.org/10.3389/fcvm.2022.864312

[33] A. Naik, G. G. Tejani, and S. J. Mousavirad, "SGO enhanced random forest and extreme gradient boosting framework for heart disease prediction," *Sci. Rep.*, vol. 15, no. 1, p. 18145, 2025. https://doi.org/10.1038/s41598-025-02525-7

[34] Chang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. A. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. *Healthcare Analytics*, *2*, 100016. https://doi.org/10.1016/j.health.2022.100016

[35] H. A. Al-Jamimi, "Early prediction of heart disease risk using extreme gradient boosting: a data-driven analysis," *Int. J. Biomed. Eng. Technol.*, vol. 45, no. 4, pp. 296-313, 2024. https://doi.org/10.1504/IJBET.2024.140563

[36] E. D. Omar *et al.*, "Comparative analysis of logistic regression, gradient boosted trees, SVM, and random forest algorithms for prediction of acute kidney injury requiring dialysis after cardiac surgery," *Int. J. Nephrol. Renovasc. Dis.*, pp. 197-204, 2024. https://doi.org/10.2147/IJNRD.S464082