FDRE TECHNICAL AND VOCATIONALT RAINING INSTITUTE

(FDRE TVTI)



የኢፌዲሪ የቴክኒክና ሙያ ስልጠና ኢንስቲትዩት
FDRE TECHNICAL & VOCATIONAL TRAINING INSTITUTE

DEPARTMENT OF INFORMATION TECHNOLOGY

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY
THESIS TITLE:

Predicting of Stunting Among Under-Two Children Using Ensemble Machine Learning Approaches

By:

Weldebirhan Tadesse Berhe

ID NO: TTMR/227/16

Principal Advisor: Hailay Beyene (Dr.)

Addis Ababa, Ethiopia
June 3, 2026

FDRE TECHNICAL AND VOCATIONAL TRAINING INSTITUTE
COLLEGE OF GRADUATE STUDIES

DEPARTMENT OF INFORMATION TECHNOLOGY

PREDICTING OF STUNTING AMONG UNDER-TWO CHILDREN USING ENSEMBLE MACHINE LEARNING APPROACHES

By:

Weldebirhan Tadesse Berhe

Student ID: TTMR/227/16

A Thesis submitted to the Department of Information Technology, College of Graduate Studies, FDRE Technical and Vocational Training Institute(FDRE TVTI), in partial fulfillment of the required for degree of Master of Science in Information Technology.

Principal Advisor: Hailay Beyene (Dr.)

Addis Ababa, Ethiopia

June 3, 2026

DECLARATION

I the undersigned, declare that this Thesis entitled "P**redicting of Stunting Among-Under-Two Children Using Ensemble Machine Learning Approaches"** is my own original work, prepared under the supervision of my advisor. It has not been submitted for the award of any diploma, degree or any other academic qualification in this or in other institution. All sources of materials used in this thesis have been duly acknowledged.

Declared by:

Name: WeldebirhanTadesse Berhe

Signature:_____

Department: Information Technology, FDRE TVTI

Date: June 3, 2026

This thesis has been submitted for examination with my approval as the candidate's advisor.

Principal Advisor: Hailay Beyene (Dr.)

Signature: 

Date: June 3, 2026

APPROVAL PAGE

This is to certify that the thesis entitled "**Predicting of Stunting among Under two children using Ensemble Machine Learning Approach"** submitted by Weldebirhan Tadesse (ID No. TTMR/227/16 in partial fulfillment of the requirement for the degree of Master of Science in Information Technology has been examined and approved by the Board of Examiners of the Department of Information Technology, FDRE Technical and Vocational Training Institute.

Examination Role	Name	Signature	Date
Principal Advisor	Hailay Beyene (Dr.)		Date: 6/3/2026
Internal Examiner	_____	_____	_____
External Examiner	_____	_____	_____
Department Head, IT	_____	_____	_____
Dean, College of Graduate Studies	_____	_____	_____

Place: Addis Ababa, Ethiopia

Date of Final Approval: _____

Acknowledgement

First and foremost extra ordinary thanks go for my Almighty God and His Mother Saint Merry, for providing me this opportunity granting me the capability to proceed successfully. This thesis appears in its current from due to the assistance and guidance of several people.

I would Therefore, like to offer my sincere thanks to all of them, Dr. Hailay Beyene, my respected Advisor, my cordial thanks for giving me the freedom to do my thesis in this area, for your warm encouragement, thoughtful guidance, critical comments, and correction of the thesis. I would like to thank you especially for friendly assistance with various problems all the time.

My sincere thanks go to the Department of Information Technology and the College of Graduate Studies of FDRE TVTI for providing the academic environment, laboratory facilities, and resources required to complete this research. I also wish to extend my sincere gratitude to all my instructors who have been the source of all my achievements. Furthermore, I would like to express my sincere gratitude to Information Technology coordinator Dr. Birknesh Weldyohans for all her enthusiasm and continuous support.

I am also grateful to Ministry of health nutrition coordination office/seqota declaration for granting permission to collect data and for the cooperation of its staff during the study.

Finally, I would like to express my special heartfelt thanks go to my parents, brothers, sisters, my wife and to my sons for their understanding and provision of unrestricted support and encouragement through both the ups and downs of my time during my study.

ABSTRACT

Child stunting is defined as low height-for-age and is a major global public health issue affecting millions of children, mainly in Sub-Saharan Africa. Chronic malnutrition, exposure to conflict and drought, and poor health services are driving up the incidence of stunting in Ethiopia. The traditional methods based on retrospective surveys are not always timely enough and do not include multifactorial. To overcome these limitations, this study develops a predictable and interpretable ensemble machine learning model to identify under-two children at risk of stunting, so enable timely interventions and resource optimization. This research utilized Secondary source of data from the Ministry of Health (Seqota Declaration), including child-level attributes, maternal-level attributes, socioeconomic, nutritional, and environmental factors from the first 1,000 days of life. Five ensemble machine learning models, such as RF, Gradient Boosting, XGBoost, Light Boost, and CatBoost were combined with advanced feature selection (Boruta), hyperparameter optimization (Optuna), and data balancing (SMOTE). Model explainability techniques, such as SHAP and LIME, were applied to provide insights into predictive mechanisms. Among the models, XGBoost achieved greater performance, with an accuracy of 81.76% and an AUC of 82.81%. Explainability analyses highlighted main factors of stunting risk, increase interpretability and policy significance. The descriptive result and analysis show that 29% (one in three) children are affected with stunted. This investigation helps to connecting methodological, contextual and age specific gaps in previous studies by integrating advanced ensemble techniques with interpretability frameworks. Finally, the prototype of this study show the threshold value, stunting status, probability of the risk (required immediate intervention or not), cause and suggested solutions. Findings are normally estimated to guide policymakers and health experts in designing targeted and evidence based interventions to decrease stunting and its long term impacts. Future work should use additional algorithms like deep learning, expand data sources and discover healthcare recommendation systems to further strengthen predictive capacity.

Keywords: Stunting, predicting, Ensemble Machine Learning, Model Explainability

Contents

DECLARATION II
APPROVAL PAGE III
Acknowledgement IV
ABSTRACT V
List of Table IX
List of Figure X
ABBREVIATION XI
CHAPTER ONE: INTRODUCTION 1
1.1. Background 1
1.2. Problem Statement 2
1.3. Objective of the Study 4
1.3.1. General Objective 4
1.3.2. Specific Objectives 4
1.4. Research Question 5
1.5. Scope 5
1.6. Significance of the Study 5
1.7. Limitations of the Study 6
1.8. Organization of the Thesis 6
CHAPTER TWO: LITERATURE REVIEW 7
2.1. Child Stunting Literature Review 7
2.2. Machine learning overview 8
2.3. Types of Machine Learning 10
2.3.1. Supervised Learning 10
2.3.2. Semi-supervised Learning 11
2.3.3. Unsupervised Learning 11
2.3.4. Reinforcement Learning 11
2.3.5. Ensemble Machine Learning 12
2.4. Hyper Parameter Optimization (HPO) 14
2.4.1. Bayesian Hyper Optimization 14

2.5. Model Explainability 15
2.5.1. Shapley Additive Explanation (SHAP) 15
2.5.2. Local Interpretability Model Agnostic Explanation (LIME) 16
2.6. Empirical literature Review 16
2.7. Research Gap Identification 22
2.8. Proposed Solution 23
2.9. Conceptual Framework 24
CHAPTER THREE: RESEARCH METHODOLOGY 25
3.1. Research Design 25
3.2. Model Architecture 26
3.3. Data Source and Collection 27
3.3.1. Data Description 27
3.4. Data Preprocessing and Preparation 30
3.4.1. Data Cleaning 30
3.4.2. Data Transformation 30
3.5. Selection of Features 31
3.6. Handling class imbalance (SMOTE) 31
3.7. Data Set Splitting 32
3.7.1. Prevention of Data Leakage 32
3.8. Model Selection and Development 32
3.9. Model Evaluation Techniques 33
3.9.1. Confusion Matrix 33
3.9.2. Performance Measure 33
3.9.3. Cross Validation Strategy for Final Model Evaluation 34
3.10. Development Tools and Computing Environments 35
CHAPTER FOUR: RESULT AND DISCUSSION 36
4.1. Overview 36
4.2. Descriptive Results and Analysis of Distribution Predictive Factors 36
4.3. Implementation Details 46
4.3.1. Handling missing values 47
4.3.2. Correlation Matrix 48
4.3.3. Dataset Splitting 49

4.3.4. SMOTE 49
4.3.5. Hyperparameter Optimization 50
4.4. Model Comparison 51
4.5. Test Performance of Models Result and Analysis 53
4.6. Feature importance 58
4.7. Prototype 59
4.8. Feature Explainability 62
4.8.1. SHAP Explanation 62
4.8.2. LIME Explanation 62
4.9. Discussion 64
4.9.1. Direct comparison with Ethiopian studies 65
CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS 66
5.1. Conclusion 66
5.2. Strengths and Limitations 66
5.3. Contributions 66
5.4. Recommendation 67
REFERENCE 68
APPENDICES 75

List of Table

Table2. 1: Empirical literature Review 20
Table3. 1: Data Description 28
Table4. 1: Hyper parameter Tuning 50
Table4. 2: The overall model performance of an experiment 52
Table4. 3: Confusion Matrix of Random Forest 54
Table4. 4: Confusion matrix of gradient boosting 55
Table4. 5: Confusion matrix of extreme gradient boosting 56
Table4. 6: Confusion matrix of light boosting 57
Table4. 7: Confusion matrix of light boosting 58
Table4. 8: Feature importance rank 59
Table4. 9: comparison to related study 65

List of Figure

Figure2. 1: Conceptual framework 24
Figure3. 1: Model Architecture 26
Figure4. 1: Stunting distribution by sex 37
Figure4. 2: stunted distributions by Child_Age 37
Figure4. 3: Stunted distributions by Maternal-education 38
Figure4. 4: stunting distribution by wealth index 39
Figure4. 5: stunting distribution by *Residence* 39
Figure4. 6: Stunting distribution by Region 40
Figure4. 7: stunting distribution by Birth weight 41
Figure4. 8: stunting distribution by Dietary_Diversity 41
Figure4. 9: stunting distribution by Drought Exposure 42
Figure4. 10: stunting distribution based on Conflict Exposure 43
Figure4. 11: stunting distribution by Diarrhoea 43
Figure4. 12: stunting distribution by Exclusive_BF 44
Figure4. 13: stunting distribution by Toilet 44
Figure4. 14: stunting distribution by Mother_Occupation 45
Figure4. 15: stunting distribution by Displacement 45
Figure4. 16: pie chart of stunting status 46
Figure4. 17: Missing values before and after Imputation 47
Figure4. 18: Correlation matrixes among attributes and stunting status 48
Figure4. 19: class distribution before and after SMOTE 50
Figure4. 20: AUC- ROC curve after HPO 52
Figure4. 21: Confusion matrix of Random forest model 54
Figure4. 22: Confusion matrix of Gradient boosting model 55
Figure4. 23: Confusion matrix of Extreme gradient boosting model 56
Figure4. 24: Confusion matrix of Light Boost Model 57
Figure4. 25: Confusion matrix of Cat Boost 58
Figure4. 26: Prototype for Not Stunted and Stunted class 61
Figure4. 27: SHAP values of Stunting Risk factors 62
Figure4. 28: LIME Prediction Explanations 63

ABBREVIATION

WHO…….World Health Organization

SSA……Sub Saharan Africa

HAZscore…….....Height for Age score

SINTA……Science and Technology Index

UNICEF……..United Nation International Children's Emergency Fund

EDHS…….Ethiopian Demographic and Health Survey

SVM……..…Support vector machine

SDGs …….Sustainable Development Goals

ML……..machine learning

SSGI……Survey Status Gizi Indonesia

SD……standard division

N……..Number of Tree

DT……...Decision Trees

RF……Random Forest

SHAPE……Shapley Adaptive Explanations

LIME…….Local Interpretable Model agnostic Explanation

RF-RFE……..Random Forest Recursive Feature Elimination

AUC……Area under the Curve

SMOTE……..syntactic minority over sampling technique

XGBoost…….....Extreme Gradient Boosting

GBoost…….....Gradient Boost

LightGBoost……..Light Gradient Boost

GOSS…….Gradient based one sided sampling

Cat Boost…….Categorical Boosting

SNNPR…….....Southern Nation Nationality Peoples' Region

TP…….True Positive

TN…….True Negative

FN…….False Negative

FP……. False Positive

TPR…….True Positive Rate

FPR……. False Positive Rate

HPO…….Hyper Parameter Optimization

ENSANUT……Mexican National Health and Nutrition Surveys

MARS…….. ……...Multivariate Adaptive Regression Splines

BDHS ……..Bangladesh Demographic and Health Survey

LR ……Logistic Regression

BMI……...Body Mass Index

ANC ……Antenatal care

CHAPTER ONE: INTRODUCTION

1.1. Background

Child stunting is characterized by poor linear growth and development, and is defined as low height-for-age. It is a significant global public health issue for children. Child Stunting is irreversible if it happens after the first 1,000 day (from the start of pregnancy to 24 months) [1] . In developing countries, child stunting results from chronic malnutrition caused by multifaceted factors such as low food intake, poor maternal and child health, low socioeconomic status, food insecurity, poor sanitation and lack of health services [2]. In 2020, 149 million children are affected with stunting worldwide, with more than 85% in Africa and Asia. In Africa, it affects about 59 million children raising the concern for strategies to struggle it [3].Although global stunting rates are improving, sub-Saharan Africa continues to struggle with higher prevalence. Eastern Africa stands out with stunting levels greater than 30% and Ethiopia is a focal point of concern. While Ethiopia has managed to improve stunting rate from 58% in 2000 to 38% in 2016, there are still regional variations [4]. Ethiopia, where war has exacerbated food insecurity, and interrupted access to health services, has high prevalence of stunting that has a serious impact on child health and development. The impact of Stunting in children under two year of age is high and it is associated with poverty. Stunting rates in low income and post conflict areas such as Tigray (48%), Afar (42%) and Amhara (42%) are too high [5].

The impacts of stunting are severe and long lasting, resulting in long-term effects on individuals and society, including irreversible physical and mental developmental delay, increased risk for chronic disease, low academic achievement and motor performance, and low productivity in adulthood. These continue in intergenerational poverty and aggravate societal inequities, and undermines national development[6]. Women and children's diet is important for children's health and survival. Good feeding practices is one of the major determinants of health, nutrition, survival and growth of children [7].In the pastoral regions of Ethiopia, cultural, knowledge, social, norm, belief, practice, household decision making and other important tasks are among the barriers leading to undermining women and children's nutrition status. Programs and strategies are suggesting identification of barriers to design interventions that improve maternal and child nutrition practices[8].

The Seqota Declaration signed by the highest government decision-makers envisages ending stunting by 2030 by a multi sectorial food systems approach. It leverages on existing nutrition programs, innovations and governance to address complex drivers of malnutrition. But to understand the dimension of stunting, it is important to work in a multi sectorial approach including health, education and social services[9].

This study provides a new perspective by examining mothers' past data and clinical data of first 1,000 day of child's life. early identification of risk factors, able to implement effective prevention mechanism to struggle stunting and improve health outcomes for children[10].Children are more affected by stunting who live in rural area than urban, children are more stunted who mother do not work, then work, children whose age 12 up to 23 months have more chance to become stunting compared to age less than 12 months, males children are highly affected compared to females and Children under two year who do not have early initiation of exclusive breast feeding, are more stunted than who have[11].

Stunting in children is occurring at first 1000 day from the start of pregnancy and is a sign of undernutrition. Children under two year of age are considered to be highly affected by malnutrition because of fast development and more prone to infectious diseases[12]. Recent development in Ensemble machine learning (ML) are providing solutions for better prediction of stunting which using a combination of algorithms and have shown better prediction power by incorporating different data resources like demography, environmental shocks and nutrition[13].Research has demonstrated the ability of ML algorithms to target high risk groups and guide intervention strategies [14].

1.2. Problem Statement

Many studies have predicted child stunting using machine learning in Ethiopia some of them use ensemble machine learning. All the researchers used national health survey data with inherent limitations, such as limited variables, focusing only on socioeconomic, maternal, nutritional and demographic variables while lack of contextual factors, miss focusing on the critical period for children under two year old, not use advanced feature selection and hyperparameter optimization. Even though in some case, ensemble machine learning models such as RF and XGBoost have highest accuracy, its lack of interpretability limits policy intervention and clinical decision making. Further, national studies lack information on household dynamics, limiting the

design and implementation of intervention strategies and mostly focus on the irreversible stunting age group [15].

In Ethiopia, stunting is prevalent, and magnified by conflict, displacement and food insecurity. Traditional approaches to predicting stunting rely on retrospective surveys and statistical models that are reactive, time-consuming, and unable to capture dynamic predictors such as climate and socio-economic shocks [16]. The main causes of stunting include malnutrition, poverty, social and cultural factors, vulnerability to infectious diseases and lack of access to health services [17].

While large-scale programs such as the Seqota Declaration exist, this calls for a need to develop interpretable, diagnostic oriented ensemble machine learning models, incorporating large and context specific data to tackle the complex causes of stunting and to support timely and targeted implementation. Despite the presentation of large scale nutrition initiatives like Seqota Declaration, stunting prevalence is still high in many regions in the country[18].

Past research had used machine, ensemble machine learning and hybrid models to predict and classify stunting of among under-five children using EDHS data set. Most of the studies do not contain program specific indicators, and not investigated true population, recent data and contextual factors (conflict exposure, drought exposure and displacement) and limited the scope to under five children, which make the study less relevant to policy, as stunting after two year of age is irreversible. Also, the national assessments are not based on the challenges of implementation of nutrition programs, particularly in the sites that are part of the Seqota Declaration intervention. Further, many of the studies focus more on predictive accuracy rather than diagnostic and model interpretability which have limited their application in policy and program design for intervention.

While prior Ethiopian studies have demonstrated the feasibility of ensemble and machine learning for stunting prediction (Ayele et al.2025 achieving 98% accuracy; Bitew et al.2022),they share critical limitation:

(1)Focus on under five children despite stunting being irreversible after 24 months, (2) use of outdated EDHS data (2011-2016) and 2019(3) absence of contextual shock variables, and (4) lack of interpretable outputs for policy makers. This study addresses each gap through age

specific targeting, recent data (2021-2026), integration of conflict/drought/displacement indicator and SHAP/LIME explainability.

1.2.1Alignment with seqota declaration

The seqota declaration, Ethiopia's flagship multisectoral nutrition initiative explicitly calls for:

- ✓ Innovative approaches to identify at risk children before stunting becomes irreversible.
- ✓ Evidence based targeting interventions
- ✓ Real time monitoring and early warning systems

This study operationalizes by:

1. Innovative approach: using ensemble machine learning with advanced tools.
2. Early identification: prediction during first 1,000 days (reversible window).
3. Evidence based targeting: feature importance ranking identifies priority intervention factors.

The model outputs (probability scores, risk thresholds, actionable causes) directly support Seqota's multi-sectorial coordination mechanisms by providing quantitative evidence for health, agriculture and social protection sectors.

1.3. Objective of the Study

1.3.1. General Objective

The main objective of this study is to identify key factors and develop a predicting of under-two child stunting by employing ensemble machine learning methods.

1.3.2. Specific Objectives

To accomplish the main objective of this study, the following specific objectives were addressed.

- ✓ To identify main features influencing children stunting.
- ✓ To select the most suitable ensemble machine learning algorithm.
- ✓ To evaluate model performance and accuracy.

1.4. Research Question

1. What are the most significant risk factors used to predict for the under-two of child stunting?

2. Which type of ensemble machine learning algorithm achieves the highest accuracy in predicting child stunting?

3. What range of performed metrics did the best-performing ensemble algorithm achieve?

1.5. Scope

The study is delimited to developing a predictive model of child stunting among young age group in Ethiopia with applied ensemble machine learning models. The study used secondary source of data from the Ethiopia ministry of health /Seqota declaration includes only the age group of under-two year. An ensemble machine learning algorithms such as Random Forest, Gradient Boosting, Extreme Gradient Boosting, Light and Categorical Boosting were used.

1.6. Significance of the Study

Stunting is a prevalent public health disorder that occurs when children practice prolonged malnutrition during critical periods of growth and development. This study is significant in several ways:

Early Identification and Prevention: Through constructing predictive models for stunting, the study able early identification of children at risk before medical signs becomes evident. This study helps for timely interventions to prevent long-term physical health and cognitive consequences of children.

Evidence-Based Decision Making: The study identifies the primary child, maternal, socioeconomic, environmental, and contextual level predictors of stunting using recent data collected by the Ministry of Health Nutrition Coordination Office under the Seqota Declaration. These results help guide evidence-based policy decisions and focused interventions.

Effective Resource Utilization: Nutrition and health initiatives can more effectively deploy resources and optimize impact by constructing predictive modeling to identify at high risk children.

Public Health Impact: Reducing stunting improves child survival, enhances cognitive development, and contributes to future productivity, thereby supporting national development goals.

Innovation in Health Research: The study constructs advanced analytical methods, including five ensemble machine learning models, which expand conventional public health research approaches. The findings can strengthen maternal and child health's programs and guide the design of more effective interventions.

1.7. Limitations of the Study

The study had significant challenges in accessed data, particularly due to the unsuitable condition. There was social awareness issues related to the importance of giving a complete history of child and mother health, including daily food information record and dataset is often complex due to the heterogeneity.

1.8. Organization of the Thesis

The study includes five chapters. The first chapter focuses on the introduction which includes background, problem of statement, objective, scope and limitation. Chapter two focuses on the discussion of literature review. Chapter three focus on over all methodology of the study, which includes detail described on the research design, model architecture, dataset, pre-processing, model development, Data balance, prediction of stunting status based on risk factors using ensemble machine learning algorithms. The fourth chapter describes the detail experimental output and analyses of risk factors of the models; explain performance measure and evaluation metrics, model results, model comparison, predictions, model explainability and prototype. The conclusion, contribution and recommendations are presented in the 5th chapter. Finally the study ends with references and appendices.

CHAPTER TWO: LITERATURE REVIEW

2.1. Child Stunting Literature Review

Stunting is defined as not growing properly because of long term deficiencies of nutrients and leads to illness. This happens when they do not eat food with nutrients and balanced diets. Stunting starts in the utero and the early year of life and builds up until the child is shorter than 97.7% of the reference population. This normally occurs at a very early age, in the first two years of life or in the first 1,000 day which is the period of growth and development. Stunting is a problem because it may occur during these critical times and it impairs the person's growth so it is important to ensure people are well nourished and don't get sick too often during these critical periods of growth. This is unfortunately irreversible because many of the physical and cognitive development processes are only one time chance and cannot be compensated later in life. While the incidence of stunting may increase after 24 months of age but the rate of increase becomes much slower beyond this period [19].

According to the World Health Organization Child Growth Standards, Stunting is a height for age Zscore (HAZ) that is measured as less than minus 2 standard deviations from the average. This means they have had long term malnutrition and growth failure over time due to insufficient nutrients and frequent infection. This means children are not healthy and strong. We can estimate whether children are stunted or not using the WHO Child Growth Standards [20].

UNICEF has a model that classifies the causes of stunting into three levels including immediate causes like not eating nutritious food and getting diarrhea often, these are reasons that affect a Childs growth. Underlying causes are issues in the family and community; these include being poor, poor water, poor toilet and hand wash, and no easy access to health care. Basic causes are reasons such as poverty, uneven distribution of wealth and political and institutional issues. This model illustrates that stunting is not just poor nutrition but also about the bigger issues in society and the economy. So to fix stunting need assistance using different approaches, from different perspective[21].Poverty and economic hardship are among the Basic Causes and most significant drivers of child stunting, as they restrict households' ability to access nutritious food, quality

healthcare, and adequate sanitation facilities. Persistent socioeconomic inequality further exacerbates vulnerability to chronic under nutrition[22].

There is household food insecurity and that is an underlying issue. Food insecurity means people do not have access to food and variety of foods. This really harms children growth. Food insecurity is bad, for children, It limits their ability to grow strong and healthy. Household food insecurity affects children's growth and nutritional status[23]. Frequent exposure to infectious diseases, such as diarrhea and respiratory infections are Immediate Causes, interferes with nutrient absorption and increases metabolic demands, so hindering normal growth and development. Inadequate dietary intake, both in quantity and quality, directly affects to chronic malnutrition and linear growth failure [24].

Maternal factors, such as low birth weight causing from inadequate maternal nutrition and poor prenatal care, are strong predictors of child stunting. Restricted fetal growth and inadequate nutrient stores at birth leads linear growth during the critical first 1,000 day of life, increasing the likelihood of chronic undernutrition [25]. Moreover, social and cultural factors affect dietary patterns, child care and health seeking practices that play an important role in children's growth [26]. In different studies the relationship between mother's age and stunting risk is different. The impact of stunting in children based the age of mother is high in less than 20 years group. The youngers mother's age increase the chance of childhood stunting [27].In among under two year of child stunting ,the lower level of mother's education have the higher chances than mothers college level and higher level of mother's education [28].

When there is Climate change, there are droughts, geographic changes, and lack of rain has an effect on nutrition. Droughts are expected to affect children greatly and lead to stunting[29].Conflict and drought have displaced people and destroyed their livelihoods. Children are starving in many parts of the country, and the number of displaced people is at wide ranging levels. In this study, conflict independent effect on stunting among children[30],[31].

2.2. Machine learning overview

Every stage of human life is now captured digitally, and the world is interconnected with data sources [32]. In current digital world, there is an abundance of data from the Internet of Things, cyber security, smart cities, businesses, smartphones, social media, health, and different sources.

The data collected from different sectors can be used to develop several smart applications in related domains. Therefore, there is an urgent need for data management technologies and procedures that can rapidly and logically find insights [33].

Machine learning is the subdivision of study which focuses on mathematical models and algorithms that computer system able to perform a specific task without being explicitly programmed. The algorithms are used for applications like data mining, image recognition and predictive modeling. The main advantage of machine learning is that procedures can function independently once trained on data [34].Machine Learning models helps in forecast the diagnosis of a disease by using historical training data[35].Numerous scientists have developed a number of machine learning models that can effectively forecast and diagnose a range of illnesses[36]. Without being specifically programmed, machines can learn pattern thanks to machine learning[37] .

Machine learning is a subsection of artificial intelligence, includes various computational models for early diagnosis and prediction such as SVM, gradient boosting, RF, and other classification models. By identifying hidden patterns in large datasets, machine learning efficiently creates predictive models. A doctor may find it challenging to predict a specific outcome because numerous factors can influence the clinical outcome. In order to discover hidden interactions between variables and outcomes and to uncover new patterns between risk factors, machine learning algorithms could more effectively take into account a variety of risk factors [38].

Computers can acquire by it and can improve due to machine learning techniques. So a computer program can learn more and improve from its past and can predict the future or make some decisions without being programmed with machine learning and past data. A more amount of labeled and semi labeled input data is used by machine learning for the model to make accurate predictions. Several diverse algorithms are used in machine learning and more ML programming languages were used in the development of these algorithms. To create a model, the algorithm is frequently given a training dataset. The machine learning algorithm use the model to predict and classify final output [33].

2.3. Types of Machine Learning

There are different types of machine learning models such as supervised, unsupervised, semi-supervised, and reinforcement learning. They have Different techniques and applicable in various real world application areas, such as in cyber security , smart cities, healthcare, e-commerce and agriculture [39].

2.3.1. Supervised Learning

Supervised learning consists of labeled data combined with target variables and the algorithm will identify the instance in the tabular data that are related to the desired output. When new inputs are presented after training, a supervised learning algorithm will identify which label the new inputs should be categorized based on the existing training data. The main function of a supervised learning model is to forecast the proper label in case of a classification problem or the estimated value of the output in case of a regression problem based on fresh input data given to the model. Supervised learning algorithms require supervision to learn the mapping between inputs and outputs and the input dataset is typically divided into training and testing dataset. The training dataset is used to train and learn algorithm, which tries to discover the underlying pattern and relationships in the data. The test dataset is used to test how well the model have learnt and how well the model is for prediction[40].

A. Random Forest (RF)

Random Forest (RF) is a supervised machine learning model used in classification and regression problem. This learning technique used more than one decision trees so, it is powerful approach, being able to deal with large data sets, noise and missing value[41]. In most of the time, RF classification approach is used to building prediction models. In many cases, when developing a prediction model, fewer variables are preferred to get a prediction because this will ease the work of data collection and make the process more efficient. There are a number of variable selection approaches for random forest classification [42]. RF is a flexible and easy to use algorithm and in some cases this technique provides good performance. RF can be created in the following steps, first it is decided how many trees will be created, and then a random sample of size N is taken. Next, a random subset of predictors is used in each tree and the second and third steps are repeated until it finish [43].

Classification accuracy of the RF model depends on the classification accuracy of each base learner and similarity between the base learner[44].

2.3.2. Semi-supervised Learning

Semi-supervised learning is type of machine learning approach that merges features of both supervised and unsupervised learning methods. This hybrid approach can be particularly useful in scenarios when more amounts of unlabeled data are readily available. In contrast to fully supervised learning, which requires a completely labeled dataset, semi-supervised learning can use the more of no tabular data to improve the model's performance, often with less labeled data. The underlying assumption is that the structure of the unlabeled data can provide constructive meaning that can be used to improve the learning process[45] .

2.3.3. Unsupervised Learning

Unsupervised learning did not have tabular data; the machine is given unlabeled data to discover the patterns. There are different types of unsupervised machine learning algorithms such as clustering, anomaly detection and association rule learning. Unsupervised learning can applied in dimensionality reduction. In machine learning algorithms like regression and classification, the number of features is often too many, to get meaning full insight the numbers of features are minimized by the process of dimensionality reduction. This effect is known as curse of dimensionality, is happen when the difficulty of modeling something is increased as the number of variables is increased. So, dimensionality reduction helps in building a model which is fast and efficient to train without losing any important information or with minimal degradation of the model[46].

2.3.4. Reinforcement Learning

Reinforcement learning is a machine learning model addresses how software agents must make decisions within an environment to enhance a concept of cumulative reward. Reinforcement learning consists of the agent making trials and errors and learning as they engage with the dynamic environment. The agent gets a reward or penalty as feedback and it modifies its behavior to maximize the reward summation with time. This ensure that the agent learns best policy, which able best action to be taken[45].

2.3.5. Ensemble Machine Learning

Ensemble machine learning model is type of supervised machine learning model, which involves training several models so-called weak classifiers by combining several machine learning algorithms. They are built upon features derived on a collection of numerous data predictions and the resulting outputs of these models are summed with various voting techniques to give better output than a single algorithm applied individually. Ensemble techniques have been highly applied to produce predictive models in a broad spectrum of clinical applications, including the diagnosis of Alzheimer disease, predicting diabetes, renal injury of severe type, spinal curvature type and type of gastric cancer [47].

Ensemble machine learning method relies on a number of issues such as the sampling, training, and combination of the base models. Various methods have been suggested of ensemble classification models. Bagging where weak classifiers are trained, on the basis of various random subsets and feature dimensions example RF. In ensemble machine learning, multiple base models are independently trained, and their outputs are combined by method, like majority voting, weighted averaging or stacking to create a more dependable and accurate final prediction. This method capitalizes on the strengths of various models and may in many case outperform any one individual model [45] .there are different types of ensemble machine learning models such as RF, AdaBoost , gradient boost, XGBoost , light boost , and cat boost are the most common [48].

A.Gradient Boosting (GBoosting)

It is type of ensemble machine learning model which depends on the boosting method to display important insight. Primarily relies on the decision trees as the underlying learner to generate a strong ensemble classifier also called gradient boosted decision tree. This consists of multiple decision tree base learners. Not all decision trees (DTs) are independent, as a new extra DT puts more weight on the misclassified samples succeeded by the previously added DTs. It can be observed that the output of previous decision tree is input to the subsequent decision. The residual is minimized with the added DT and the loss decreases along the negative gradient direction with in iteration. finally output of the forecasting is the end result of the addition of results of all DTs[49].

B.Extreme Gradient Boosting (XGBoosting)

This is type of ensemble machine learning use Gradient boosting algorithm. XGBoosting is a powerful algorithm used in the application of regression and classification. XGBoost has been a recent winner of the applied machine learning. XGBoost is an algorithm which uses maximum tree depth, learning rate and subsampling to avoid overfitting[48]. XGBoost, is not susceptible to overfitting [50]. XGBoost algorithm needs minimum feature manipulation such as normalization and scaling also controls such cases, besides, it can handle missing values. This algorithm gives an importance of features and this enables understanding of input features as well as the selection of features. XGBoost is fastest than other type ML algorithms, it is able to process large data volumes, and it is not vulnerable to overfitting. Moreover, it is often more effective than other ML algorithms[51],[52].

C. Light boosting (Light Boost)

It is a well-organized application of the gradient boosting algorithm. It can be applied to classification, ranking and other types of machine learning. The algorithm Light Boost uses two new procedures; Gradient based one sided sampling and Exclusive Feature bundling which makes much faster and more accurate to train. A variation of gradient boosting, the GOSS technique focuses more on training instances with higher gradients, which leads to faster learning and reduced computational complexity[48].Light Gradient Boosted is better than other algorithms like, it has a shorter training time and more efficient; it uses discrete values instead of continuous values, it is more accurate, and lastly it works better with large datasets[53].

D. Categorical boost (catboost)

CatBoost algorithm is a gradient boosting algorithm [54].In the process of training, categorical features are effectively allocated with by the algorithm. One thing that has enhanced with CatBoost is the unbiased estimation of gradients, which minimizes over fitting[55]. The other significant of Cat Boost model is that automatically transforms categorical variables into machine learning understandable format of numbers. Categorical features have distinct set of values which are described as categories and are largely not comparable. The most frequently used type of encoding is one-hot encoding, which is used on low-cardinality categorical features, where the original feature is substituted by a binary one[48].Cat Boost utilizes a more effective

and robust method that prevents overfitting and guarantees that every sample in the training dataset contributes to the model [56].

2.4. Hyper Parameter Optimization (HPO)

HPO is the method of finding the best combination of hyperparameters for a given machine learning model to increase the result of particular task. Hyperparameters are parameters that are set before the training process established and cannot be learned directly from the data, unlike the model parameters that are learned during the training process[57]. Optimization has been an integral part of a decision making process. It is a process of finding the best possible solution of a given problem[58] . Hyperparameter optimization is significant and widely used technique in model development. To Enhance model output, selection of hyperparameters contributes important role [59] .

Hyperparameters that prevent overfitting can regulate a model's complexity and its capability to simplify. Regularization, hyperparameters enables to protect overfitting with monitoring the capacity of the model. The hyperparameter tuning can support to handle a balance between the model fit and overfitting to get better performance on output. Hyperparameters reduced training time, such as the number of training looping, can significantly impact the training time of a model. By optimizing these hyperparameters, you can making the model development process more efficient [60]. There are different types of hyperparameter optimization techniques, such as grid , random search, optuna Bayesian optimization, Hyperband, and racing[61].

2.4.1. Bayesian Hyper Optimization

Bayesian Hyper Optimization is a strong optimization method that is used to optimize expensive functions, such as the hyperparameters of the models. Bayesian optimization is a more efficient hyper parameter optimization technique. Bayesian optimization is more appropriate for complex relationship factors, non-convex and noisy optimization problems. In contrast to outdated optimization techniques that need numerous evaluations, Bayesian Optimization excels in addressing costly, noisy functions[62].

Optuna Bayesian optimization

Optuna is best tunes hyperparameters, quickly improving model performance. It utilizes a mixed of Bayesian optimization and complex procedures to recognize the optimal hyper parameter configurations to specific model. It operates by developing a surrogate model able to guess the effectiveness of different hyperparameter configurations and subsequently employing the model to guide the search for optimal parameters. The machine learning Hyperparameters optimization has been shown to be superior in terms of effectiveness and performance in optimization and enhancing final model output[61].

Some research was used grid and random search to apply hypermeters, but those methods, are not efficient because of randomly observe the hyperparameter space, resulting in increased processing costs, used only in the case of a low dimensional hyper-parameter space and have very low results, especially for complicated models such as gradient boost and extreme gradient boost. Whereas, Optuna's Bayesian hyperparameter optimization allows to search for best hyperparameters more intelligently and get faster convergence and more accurate models [63],[64].

2.5. Model Explainability

Model explainability is the ability to define and understand the decisions and predictions produced in machine learning model. It seeks to offer information on how the model came to its outputs to enhance transparency, trust and accountability in the decision making process. Explainability in models is critical in areas where the implications of erroneous or biased forecasts may be very important, including healthcare, finance and legal systems. Model explainability is a collection of methods used to help users or developers of machine learning (ML) models to understand the reasons behind the behavior of the models. Explanations may be of various types, e.g., informing patients about the symptoms that predetermined a certain diagnosis or helping factory workers to analyze bottlenecks of a production process[65].

2.5.1. Shapley Additive Explanation (SHAP)

A Shapley Additive explanation is a technique to interpret the predictions of machine learning models. SHAP values offer a consistent framework of variable attribution and interpretability of global models. SHAP values have their foundation in the cooperative game theory and are

inspired by Shapley values. It gives a value to every feature in a prediction, which implies the contribution to the prediction result. SHAP method computes the Shapley values based on the model by considering different mixtures of input variables and computing the difference in result of prediction between the presence and absence of a feature. This method enables one to comprehensively understand how individual features contribute to shaping of model predictions[66].

2.5.2. Local Interpretability Model Agnostic Explanation (LIME)

In machine learning model, Local interpretability Model agnostic Explanation is a technique for interpretability and explanation. It is a model-agnostic method that aims to provide local descriptions for individual predictions made by a trained model. LIME is a technique used to examine and discover the connection between input feature and the outcome by a pre-trained model. It is a model-agnostic approach that focuses on local interpretability, examining the impact of adjusting input parameters on the resulting output. Modifying the inputs and observing the resulting changes, LIME provides insights into the accuracy of model predictions and finds which input factor affected the prediction of an exact data model[66].

When physicians purpose to check if a young child is developing properly, it's crucial for them to understand how they form their assessments. They apply exact tools such as SHAP and LIME to explain what features, such as the child's nutrition, the mother's emotional state, and the home atmosphere is important for the child's development. This helps everyone trust the answers and come up with the best ways to help Child grow up healthy and strong[67].

2.6. Empirical literature Review

For investigating child stunting of under-two using ensemble machine learning algorithm, I reviewed the latest and recent scientific papers for the purpose of identifying drawback and to fill those missed one with applied ensemble machine learning model.

Chilyabanyama et al. (2022) used 2018 Zambia DHS to train and compare LR, RF, SVM, XGBoost, and Naïve Bayes models. Random Forest has the higher accuracy of 61.6%. Feature selection with Random Forest selects 13 basic predictors like age of children, sex, and mother educational level and so on. A limitation was the lack of interpretability and reliance on a dataset not originally tailored for predictive modeling[68].

Pratama et al. (2024) compared several machine learning regression models to forecast stunting prevalence across Indonesian provinces using national data from sources including SSGI and the Ministry of Health. Random Forest Regression achieved the highest performance with 90.54% accuracy. Although highly effective, the study focused on regression rather than individual-level classification and highlighted the need for more granular/not smooth, child-level data for intervention planning[69].

Shen et al. (2023) used several ML models, such as LR, decision trees, SVM, and XGBoosting to explore data from Papua New Guinea DHS conducted from 2016-2018 including 3380 children. Feature selection was performed using LASSO and RF-RFE. XGBoosting model achieve highest predictive efficacy AUC of 76% and accuracy of 72.8%.Key predictors included geographic region, household wealth, and child's age and birth size. The study highlighted the power of feature selection but noted model complexity and lack of interpretability[70].

Ndagijimana et al. (2023) examined Rwanda, utilizing the DHS dataset from 2019 to 2020. Several ML models were assessed and Gradient Boosting model yielding the best performance with accuracy of 79.3% and AUC of 89%.Important features included mother height, child age, residence, and access to television. While the model displayed high performance, the authors distinguished the restricted availability of domain specific features and potential challenges in generalizing the model across different places. The limitation of this study are Data imbalance and limited domain specific features[2].

Fogarty et al. (2025) examined stunting in Mexico through a substantial dataset from the Mexican National Health and Nutrition Surveys (ENSANUT) conducted in 2006, 2012, and 2018, which included 20,803 children aged fewer than 5. They evaluated several algorithms, including XGBoost, RF, Neural Networks, Multivariate Adaptive Regression Splines (MARS), and Light Gradient Boosting, with XGBoost achieving the highest Youden Index (0.255). The primary key indicators consist of socioeconomic status, residential location, child's age, indigenous identity, family structure, and local deprivation metrics. Together, these studies emphasize the usefulness of machine learning in forecasting child stunting in various settings while pointing out a shared deficiency in converting predictive models into diagnostic tools for real-world health initiatives[71].

Rahman et al. (2021) examined the main factors and forecasted stunting in children under 5 years in Bangladesh use Bangladesh Demographic and Health survey (BDHS) from 2017 up to 2018 dataset, which included 8,000 children. They utilized LR for detection and SVM, RF, and LR for prediction, with RF achieve best accuracy of 88.3% in forecasting stunting. Nonetheless, the predictive emphasis of the study constrained its range to particular risk factors such as BMI, ANC visits, vitamin A supplementation, and breastfeeding duration that were not included[72].

Zemariam et al. (2025) effectively applied machine learning to predict stunting among adolescent girls using EDHS data, their study is limited by its focus on a non-priority age group for stunting prevention, dependence on a general purpose dataset lacking intervention and food security indicators, and absence of program specific contextual analysis such as the Seqota Declaration. Furthermore, limited model explainability restricts translation into actionable nutrition policy. These limitations shows a clear population, empirical, contextual, methodological, and application gap [73].

Hendy et al. (2005–2014) applied supervised ML to predict stunting among under five Egyptian children using Egyptian DHS data. RF and Gradient Boosting achieved accuracy of 90% and ROC-AUC of 96% respectively, outperforming Logistic Regression, XGBoost, and KNN. Key predictors included maternal education, child's nutritional status, birth size, wealth index, and residence. The study demonstrated ML's strong potential for early identification of vulnerable groups, though limited by outdated datasets[74].

Ayele et al. used EDHS data (2011–2016) to build a multi-class ensemble ML model predicting stunting severity. After applied preprocessing and data balancing, Random Forest has high accuracy of 98% and ROC-AUC of 99.99, but suffers from low interpretability. The main factors are child age, maternal education, birth order, wealth index, BMI, breastfeeding duration, and sanitation access. The study advanced beyond binary classification, offering granular insights for targeted interventions but missing contextual factors such as conflict exposure, drought exposure ,climate shocks, limited model explainability and not use advanced feature selection and hyperparameter optimization techniques [15].

Wudu et al. used EDHS data (2019) to build a multi-class ensemble ML model predicting stunting severity. After applied preprocessing and data balancing, hybrid achieved high accuracy

of 94%, improve performance, but suffers from global low interpretability and accuracy alone is insufficient. Important predictors included child's age, birth interval, and region. The study advanced beyond binary classification, but missing contextual factors and limited global model explainability, limited model selection for categorical variables and not use advanced feature selection and hyperparameter optimization techniques[75].

Table2. 1: Empirical literature Review

Authors(years)	Problem investigated	Datasets	Methodologies	Country	Best model accuracy	key Gap
Chilyabanyama et al, 2022	Evaluate ML classifiers for predicting stunting in Zambia	ZDHS 2018 and data 6799 instances	RF, XGBoost, SVM, LR, Naïve Bayes	Zambia	RF 61.6%	Low interpretability
Pratama et al, 2024	Compare machine learning regression models for stunting prevalence	Indonesian national data (SSGI)	SVR, Decision Tree, RF Regression, GAM	Indonesia	RF Regression 90.54%	Nation level data only, lacks individual level granularity and not design for prediction
Shen et al, 2023	Identify optimal ML model for stunting prediction in PNG	PNG DHS 2016–2018 and data 3380 instance	LASSO, RF-RFE, XGBoost, SVM, Decision Tree, LR	Papua New Guinea	XGBoost 72.8% and AUC 0.765	Complex models, small data, lack of interpretability, population overlook and computational cost
Ndagijimana et al, 2023	Build predictive stunting model using ML in Rwanda	RDHS 2019–2020 and data 3814 instance	GBM, RF, SVM, NB, LR	Rwanda	GBM 79.3% and AUC of 0.89	Data imbalance, limited domain specific features
Rahman et al, 2021	Examine risk factors and predict stunting in children	BDHS 2017-2018 and data 7079 instance	LR for detection, SVM, RF, LR for prediction	Bangladesh	RF 88.3%	limited to specific risk factors and dataset, lack of used advanced methodology and population overlook
Zemariam et al, 2025	predict stunting and identify key determinants among adolescent girls in Ethiopia using machine learning algorithm	EDHS 2016,sample of 3,156 instance with adolescent girls (15–19 years)	Smote Random Forest, XGBoost, SVM,	Ethiopia	RF 77% and AUC of 85%,	non-priority age group, Empirical/Data Gap, Contextual Gap, Methodological Gap ,e and use outdated data

Fogarty et al,2025	Understand social economic predictors of under-5 stunting	Mexico ENSANUT 2006, 2012, 2018 and n=20,803	XGBoost, RF, NN, MARS, LGB and Logistic Regression	Mexico	XGB achieved highest Youden Index (0.255)	population overlook , limited maternal/fetal data
Hendy *et al*, 2025	classification and prediction of stunting among under-five Egyptian children	Egypt (DHS) conducted in 2005, 2008, and 2014	Grid search , Random search, SMOTE,XGB, LR, RF, GB, and K-NN	Egypt	GB and RF exceeding 90% and ROC-AUC above 0.96	Did not use recent data, population overlook , Contextual Gap, Dataset Gap,
Ayele,2025	Predicting stunting status among under five children in Ethiopia using ensemble machine learning algorithms	EDHS data from 2011 to 2016, 18,451 instances with 28 features	Grid search ,Random search, SMOTE, ,RF, AdaBoost, XGB, and CatBoost	Ethiopia	RF 97.985%	population overlook(Population-Specific Gap) , Contextual Gap, Dataset Gap(out dated) ,methodological gap(Boruta, optuna and SHAP and LIME)
Wudu et al,2026	Explainable hybrid model for predicting stunting and identifying key risk factors among Ethiopia children under five	EDHS data 2019,11121 instance with 21 features	Grid and Radom search, LIME,RF, XGB, Extra tree, MLP and Hybrid	Ethiopia	Hybrid 94%	Population overlooking, contextual gap, methodological gap , dataset gap , model selection gap for categorical variable in health sector

2.7. Research Gap Identification

In Ethiopia a number of studies have used machine learning and ensemble models such as Random Forest, XGBoost, and adaptive Boosting to forecast child stunting using data from national health surveys or the DHS, those researchers that have common limitations such as:

1. Lack of Interpretability

A lot of models that perform well such as Random Forest and XGBoost, but reduces their usefulness for targeted interventions and policy making. Because the researcher did not focus Interpretability issue for policy level insight using Global and Local Interpretability such as SHAP and LIME to showing why a specific child is predicted as stunted and did not seen individual child-level explanations to guide community health workers.

2. Limitations of the Dataset and Data Integration Gap

Predictive modeling was not a specific purpose of the DHS and national survey datasets that were employed. Moreover, features that is important for predicting stunting, such as drought, conflicts, displacement, and food aid, place of delivery are frequently absent.

3. Geographic Gap

Studies focused on forecasts at the national level, which makes it more difficult to create useful insights at the urban vs rural and regional disparities level where interventions are most needed.

4. Methodological Gap

In Ethiopia, Prior stunting prediction studies were used wrong Feature Selection methods such as filter and wrapper methods to select attributes with linear relationship, but filter and wrapper methods are not suitable for attributes having nonlinear relationship. To choose the best predictor they did not have to consider how different feature selection methods affected ensemble model performance, generalizability, and interpretability because stunting uses attributes having with complex relationship. Boruta method of feature selection is best to solve this problem since Boruta Captures nonlinearities and is robust to noise. Even though gird and random searches were used, but Optuna is best hyper parameters optimization to stunting.

5. Age-Specific Gap

Prior research focused on under-five age group of children. Focused on prediction for under-two children is rare, despite this being the most critical window for stunting prevention. Most studies focused on under five children, overlooking the critical under-two window when stunting prevention is most effective and irreversible damage occurs.

2.8. Proposed Solution

The proposed solution addresses the identified gaps through an integrated technical framework:

Gap 1: Lack of interpretability solution

- ✓ SHAP provides global feature importance, which factors matter most nationally.
- ✓ LIME local explanations explain why a specific child is at risk.
- ✓ Enable policymakers to understand factors; enabling health workers to counsel families.

Gap 2: Dataset limitation solution: seqota declaration data (2021-2026)

- ✓ Includes contextual variables absent in DHS
- ✓ Recent data captures current conditions
- ✓ Program specific indicators align with intervention targeting

Gap 3: Geographic/ regional gap solution: regional level prediction

- ✓ Model includes region as categorical features 13 Ethiopian region.
- ✓ Enables stratification for targeted resource allocation.
- ✓ Identifies hotspots of Amhara, Afar, and Tigray with highest predicted risks.

Gap 4: Methodological gap solution: Advanced techniques

- ✓ Boruta for feature selection (handles non-linear relationship, robust to noise)
- ✓ Optuna for hyperparameter optimization (efficient, avoids local optima)

Gap 5: Age specific gap solution: Under-two focus

- ✓ Critical window for prevention (stunting is reversible before 24 months)
- ✓ Enables early intervention rather than late compensation
- ✓ Aligned with WHO/UNICEF first 1000 days framework

2.9. Conceptual Framework

The framework uses an ensemble machine learning approach to analyze how various independent variables affect child stunting, a binary outcome (dependent variable) defined as stunted or not stunted. These factors are processed by an ensemble machine learning models, such as RF, gradient boosting, XGBoosting, Light boosting and cat boosting which predict the likelihood of stunting and indicate for intervention based on the risk levels with different sectors. The workflow starts with inputting independent variables into the ensemble machine learning model, which trains and predicts stunting outcomes, ultimately report to policy recommendations to address stunting effectively.

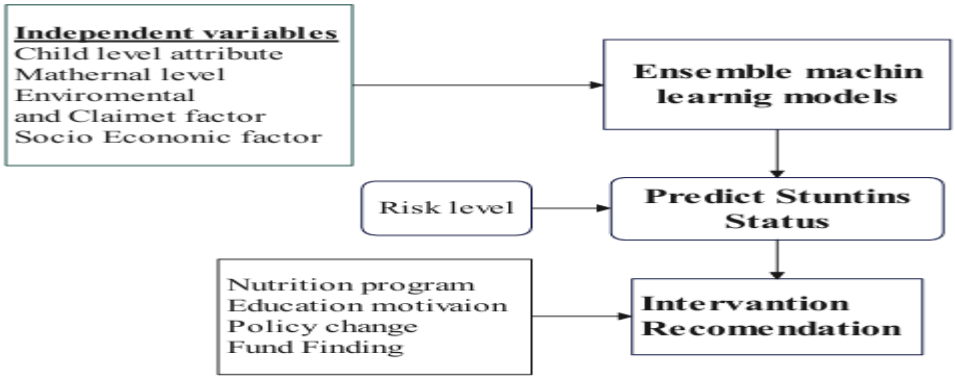


Figure2. 1: Conceptual framework

CHAPTER THREE: RESEARCH METHODOLOGY

Research methodology outlines systematic approach to develop interpretable, predicted and diagnostic focused on ensemble machine learning models to predict under-two child stunting using ensemble machine learning models, to address gaps in interpretability, dataset limitations, diagnostic focus, aggregated analysis, and underutilization of advanced techniques. The following subsections contain the detailed ordered steps of the methodology that I include in my study.

3.1. Research Design

A research design is the organized collection and analysis of data, along with a working scheme, in a manner that aims to overcome the research problem. This study is initiated to come up with a model for the risk factor prediction of stunting. For this purpose, the research employs an experimental design, which involves the execution of an algorithm and assessing the efficiency and effectiveness of those methods through experimental evaluation through addressing practical issues. Experimental research design involves dataset preparation, preprocessing, feature selection, data splitting, data balancing, model building, validation, performance evaluation and interpretation. Extensive experiments were conducted following a structured procedure. Therefore, the study has examined the feature selection process, which is suitable for identifying the risk factors for Stunting. The data preprocessing tasks, including handling missing value, outlier detection, and feature engineering, have been carefully undertaken. Additionally, the classification techniques have been intensively assessed to determine the best and suitable model for predicting the risk factors of stunting. The experimental research design approach allows us to establish a causal relationship between the risk factors and the target variable by systematically manipulating the independent variables and observing their effects on the outcome.

The study utilizes secondary data from the Seqota Declaration monitoring dataset, a nationally coordinated, multi-sectorial initiative in Ethiopia aimed at ending child stunting. The dataset integrates health, nutrition, socioeconomic, environmental variables, providing a comprehensive basis for stunting prediction.

A correlational approach is employed to examine associations between child stunting and explanatory variables related to child, maternal, household, and environmental characteristics. Establish a causal relationship with identify associations and predictive patterns. An ensemble machine learning driven predictive design is applied using different ensemble learning algorithms, with an emphasis on model performance and interpretability through feature explainable techniques such as SHAP and LIME to support diagnostic and policy relevant insights.

3.2. Model Architecture

To address the gaps that have identified in child stunting prediction studied by different researcher new model architecture is proposed. This architectures object is to develop an interpretable, diagnostic focused machine learning model for early child stunting detection and prediction.

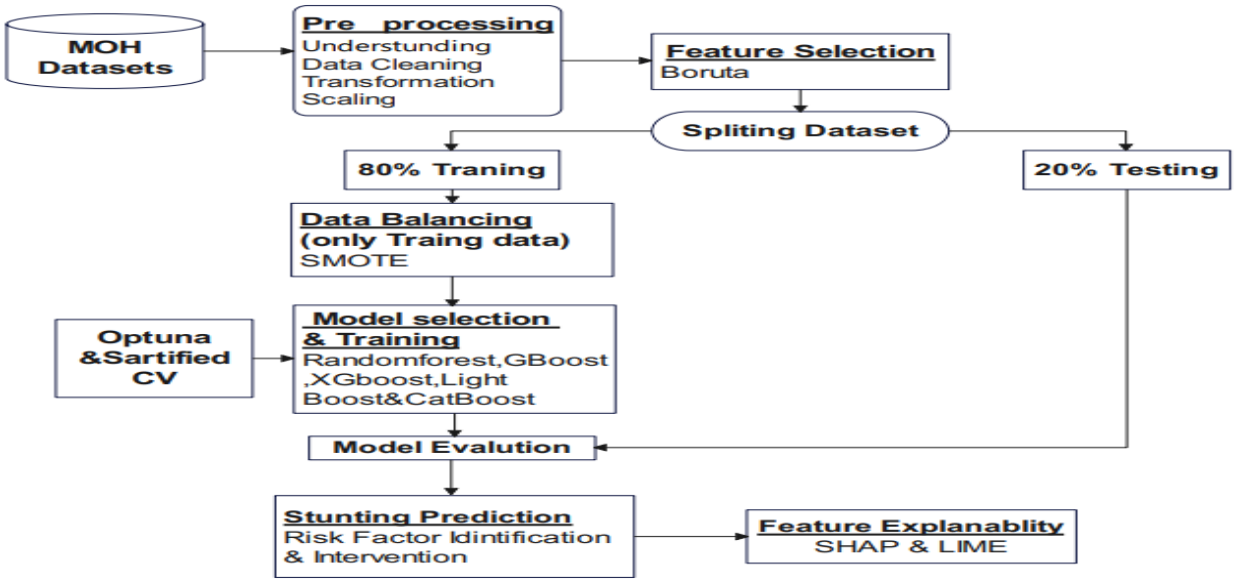


Figure3. 1: Model Architecture

The above Figure illustrates the general approach of the risk factor child stunting prediction model. This model has seven major components. The first step involves input the collected data and the second step preparing the data using different preprocessing systems, such as data understanding, cleaning and transformation. The third step involves feature selection. The fourth section involves data set splitting into training and testing sets. The fifth section involves data

balancing for only training data. The six steps involve for model selection by training and validating the machine learning models applying hyper parameter optimization techniques such as optuna to models. The seven steps involve prototype prediction based on the best model. Finally, models interpretability is applied for the best model.

3.3. Data Source and Collection

A variety of data sources would have been employed in the conceptual study of child stunting under-two years old in order to create a general predictive model. The study used secondary data source from Ministry of Health Nutrition Coordination Office with Seqota Declaration gives important information about factors of child stunting.

3.3.1. Data Description

The data description is a complete description of a data set. It includes details of the data source, columns in the data set and other key information necessary to interpret and analyses the data. This description helps users understand the data and its possible uses before working with it. For this study, data was gathered from Ethiopia Ministry of health, seqota declaration. The dataset contains 27 attributes including the target variable. Based on the following Table 3.1 dataset description contains 8 numerical data and 19 categorical data such as Sex 0 for Female and 1 for Male, Diarrhoea 0 for No and 1 for Yes , Exclusive Breast feeding 0 for No and 1 for Yes, Maternal Education 0 for No Education ,1 for Primary ,2 for Secondary and 3 for Higher, Maternal Anaemia 0 for No and 1 for Yes, Mother Occupation 0 for No and 1 for Yes, Marital Status 0 for Single,1 for Married ,and 2 for Divorced, Residence 0 for Urban and 1 for Rural, Wealth Index 0 for Poor , 1 for Middle and 2 for Rich, water 0 for Unimproved and 1 for Improved , Toilet 0 for Improved and 1 for Unimproved , Place of Delivery 0 for Home,1 for Hospital and 2 for Other, Region 0 for Addis Ababa ,1 for Afar,2 for Sidama,3 for Amhara, 4 for Benishangul-Gumuz ,5 for DireDawa,6 for Gambella,7 for Harari ,8 for Oromia ,9 for SNNPR,10 for Somali,11 for South West and 12 for Tigray, Market Access 0 for No,1 for Yes, Drought Exposure 0 for No,1 for Yes, Conflict Exposure 0 for No,1 for Yes, Displacement 0 for No,1 for Yes and Food Aid 0 for No,1 for Yes those are all Categorical.

Table3. 1: Data Description

No	Attributes	Type	Description	Categories
1.	Child Age	Numeric	What is an age of the children in month?	In month(0-23)
2.	Sex	Categorical	What is the sex of the children?	Female or Male
3.	Birth weight	Continuous	What is weight of the children at birth in kg?	1.4, 1.5,1,6…
4.	Diarrhoea	Nominal	Is the children had Diarrhoea in the last 2 weak?	No or yes
5.	Exclusive-BF	Nominal	Is the child breast feed for the first six months?	No or yes
6.	Mother Age	Numeric	What is the age mother in (year)?	14, 15…
7.	Mother Education	Categorical	What is the educational level of mother?	No-education, Primary, Secondary or Higher
8.	Mother Anaemia	Nominal	What is Anaemia status of the mother?	Not Anaemic or Anaemia
9.	Mother occupation	Nominal	Is that Mother currently working?	No or yes
10.	Marital status	Categorical	What is the marital status of Mother's?	Single, Married or Divorced
11.	Mother height	Continuous	What is mother's height in centimeters?	12, 13, 14…..
12.	Mothers BMI	Continuous	What is Mother's body mass index?	12….

13.	Antenatal Visit	Numeric	How many time mother visits during pregnancy?	0, 1, 2….
14.	Residence	Categorical	Where is residence place of the children?	Rural or Urban
15.	Wealth index	Categorical	What type of wealth index of the child family?	Poor, Middle or Rich
16.	dietary diversity	Categorical	How much dietary diversity taken daily?	0-7
17.	Water	Categorical	What type of water the family members have?	Unimproved or improved
18.	Toilet	Categorical	What type of toilet the family members have?	Unimproved or improved
19.	Delivery Place	Categorical	Where is the birth place of the children?	Home, health or Other
20.	Region	Categorical	In where Region of Ethiopia is the children live?	Addis Ababa, Afar, Sidama, Amhara, Harari, Oromia, SNNPR… and Tigray etc
21.	Market access	Nominal	Is that having market access distribution?	No or Yes
22.	Drought Exposure	Nominal	Is that the place Exposure to drought?	No or Yes
23.	Conflict Exposure	Nominal	Is there having armed conflict?	No or Yes
24.	Displacement	Categorical	Type of Displacement?	No, internal or external
25.	Food Aid Received	Nominal	Is there emergency food aid supported?	No or Yes
26.	Stunting Status	Categorical	What is the Stunting status of children?	Not Stunted, Stunted

3.4. Data Preprocessing and Preparation

Data preprocessing and preparation included cleaning, and transformation (missing value, noise and detecting outliers), data transformation (encoding, discretization, and normalization) and feature selection were performed. The primary goal of preprocessing is to acquire good quality data [76].

3.4.1. Data Cleaning

Data cleaning is the process of identifying missing values and incorrect data in the dataset. Several data cleaning approaches have been established to address these issues by filling in missing values, noisy data, discard outliers, and resolving inconsistencies. Data cleaning is critical because incorrect data creates problems in understanding and finding of insights [77].

The general purpose of data cleaning is to control and manage missing values, noise and outliers.it can solve the problem of missing value using imputation problem solving techniques such as mean and mode. Along with the missing value, the data will also have noise, outlier and inconsistent values for different features. Hence, in this case, it would have use outlier detection techniques[78].

3.4.2. Data Transformation

Data transformation is the process of translating from one cleaned data format to other for the propose of understanding by machine learning .there are different types of data transformation such as one hot encoding, labeled and ordinary encoding. Encoding is the process of converting categorical features to numerical values Encoding converts categorical variables into numerical values. The selection of encoding technique depends on the nature of categorical features, types of problem and the selection machine learning model. One Hot Encoding is the most encoding method applied in converting categorical features to machine learning format of binary values of 0 and 1.the second type of encoding technique is Label encoding, which label unique numbers to each of different features in range of numbers [79].Ordinal Encoding is also type of encoding technique it is similar with Label except it encodes based on the hierarchy of the categorical feature, it assign values like 0,1,2,3[80]. Another type of transformation technique is Normalization or scaling, which involves making range of continues numeric values by categorized into similar scale of categorical feature. Some features may dominate the model due

to large numeric ranges, Adjusts the range of numerical values so that features are on a similar scale is known as Normalization or Scaling of Numerical Features [81].

3.5. Selection of Features

The process of choosing highly significant features based on benchmark is known as feature selection. The aim of feature selection is to pick the features in a data set that are useful and drop other as irrelevant or unnecessary. An organized method is used to determine which features are most relevant and significant when selecting and arranging them for inclusion in the predictive models this process would guarantee a well informed and evidence based feature selection that will significantly improve the accuracy levels of prediction. The study used Feature Selection Tools for eliminating irrelevant predictors and to Identifies key factors influencing stunting[82].

The study used Boruta feature selection technique because of its capacity to manage high dimensional datasets and find statistically important features. The Boruta method, a variation of the Random Forest algorithm, evaluates the relevance of each feature using statistical testing and shadow characteristics. Boruta feature selection selects best important and irrelevant features by evaluating the significance of each original feature in relation to its corresponding shadow feature. Using Boruta on these datasets, it is possible to find important parameters linked to disease of predictive models[83].

3.6. Handling class imbalance (SMOTE)

Synthetic minority oversampling technique (SMOTE) is a method for oversampling instances dealing with imbalanced classification categories in dataset. Many real world data model applications involve learning from imbalanced data sets, where the particular events of interest may be very few when compared to the other classes. If the dataset has class imbalance, study applied data sampling strategy particularly SMOTE. A challenge performing statistical inference and modelling on imbalanced datasets is that the implications and outcomes from such analyses will influence the majority class. Other solutions under sample the majority class to be equivalently represented in the data with the minority class. SMOTE in the models, oversamples the minority class ,Rather than under sampling the majority class[84].

3.7. Data Set Splitting

Dataset splitting is process of dividing the total data set in to training and testing set to discard bias and to avoid over fitting to training data in Machine Learning Models. This is always the practice of data analysts and data scientists to avoid the machine learning algorithms to produce an overfitting type which may not perform well in the real test data, data is divided into series of train sets and test sets in order to train different parameters. In research use 50:50, 60:40, 70:30, 80:20, and 90:10 train test ratios for the splits. By using train and testing sets, the train set will be used to train the model to learn the inputs that are used to get good result; the test set will be used to test the effectiveness of the model [85].

3.7.1. Prevention of Data Leakage

Data leakage happens when data from the test set is used in training process resulting highly optimistic performance estimates. This study implements multiple precautions:

1. Temporal leakage prevention

- ✓ All data collected prospectively (2021-2026)
- ✓ No feature information used to predict past outcomes.
- ✓ Features are those available at time of prediction (no post outcome measurements).

2. Preprocessing leakage prevention

- ✓ Missing value imputation: using mean/mode calculated only on training set then applied to test set.
- ✓ Feature scaling: fit scaler on training set only.
- ✓ SMOTE: after splitting the data must be applied only to training set, never to test set.
- ✓ Boruta feature selection: performed on training set only.

3. Validation leakage prevention

- ✓ Cross validation folds maintain class proportion.
- ✓ No information from validation fold used in training (including imputation statistics).
- ✓ Final test set exactly once.

4. Variable leakage check

- ✓ Verify no future variable (subsequent child health outcomes known at prediction time).
- ✓ Confirmed outcome variables (stunting status based on HAZ at 24 months) not used as predictor because it create model overfitting

3.8. Model Selection and Development

In this study, a comprehensive ensemble machine learning approach was implemented. By incorporating a set of ensemble methods, including Random Forest to Robust against over

fitting, great for feature selection, Gradient Boosting & XGBoost for High predictive accuracy for complex, non-linear interactions, CatBoost & Light Boost to Faster computation for large datasets, efficient feature engineering. Ensemble machine learning approaches mix several models to increases the robustness of the predictive model and enables full understanding and exact forecasting of the pattern.

3.9. Model Evaluation Techniques

Evaluation metrics allow for measuring the performance of machine learning models. This study would have used model evaluation metrics to assess the performance predictive model for child stunting such as accuracy, recall, precision, F1_score, and ROC curve.

3.9.1. Confusion Matrix

Confusion Matrix is considered a basic and easy to understand metric for measuring the accuracy and correctness of a model. It consists of values that represent both correct and incorrect classifications and is commonly applied in the evaluation of different classification models. This matrix enables a comparison of the number of data points that are accurately versus inaccurately classified or predicted, making it particularly useful for classification scenarios involving multiple classes. True Positive (TP) indicates that the numbers of children are actually stunted and the model predicted as stunted. True Negative (TN) shows that the numbers of children are actually not stunted and the model correctly predicted as not stunted. False Positive (FP) show that the children are actually not stunted, but the model incorrectly predicted as stunted. False Negative (FN): shows that the numbers of children are actually stunted, but the model mistakenly predicted not stunted or missed stunted case so that this is the worst predication case.

3.9.2. Performance Measure

Accuracy is used to know the efficiency of models which reflects the occurrence of correct classifications of data points.

- ✓ **Accuracy:** used to calculate the degree of correctly predicted value out of the total data set. Accuracy is calculated dividing the sum of TP and TN by the sum of TP, TN, FP and FN.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.1)$$

- ✓ **Precision:** representing the quality of accurate predictions produced by the model, serves

as a metric to measure its performance. It is actually calculated by dividing the sum of all positive predictions by the ratio of actual positives such as both TP and FP.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.2)$$

- ✓ **Recall:** referred to as the True positive rate, measures the machine learning model's capability to correctly recognize a high proportion of actual positive instances. It highlights the model's efficiency in capturing positive cases within the dataset. Recall computes the competence of a model to remove missing value of actual positives, minimize FN rate and checking a complete identification of positive instances within the models dataset.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.3)$$

- ✓ **F1-score:** is the basic matrix that measure the accuracy of machine learning model by balancing recall and precision and it is can calculate as harmonic average of Recall and Precision. The F1_score, which is obtained by taking the average of precision and recall, gives equal weight to recall and precision.

$$\text{F1_score} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.4)$$

- ✓ **ROC**: is a performance metric mostly used in classification tasks. In the area under the curve (AUC), a higher True positive rate (TPR) ratio is preferred than lower false positive rate (FPR). The value of AUC is range 0.5 to 1 and the higher value is the best.

$$ROC = \frac{TP + TN}{TP + TN + FP + FN} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3.5)$$

3.9.3. Cross Validation Strategy for Final Model Evaluation

Whereas Optuna employed 5-folds cross validation for hyperparameter optimization, the final model evaluation uses separate 10-folds stratified cross validation to confirm strength and avoid overfitting.

Processes

1st The training set 12,188 after SMOTE is divided in to 10 equal folds, used to maintaining class proportion (stratification).2nd For each fold i start from 1 to 10, train on 9 fold (90% of training

data) and validation on 1 fold (10% of training data).3rd Used Performance metrics such as accuracy, precision, recall, f1-score and AUC. As a final point, the model is retrained on the whole training data and evaluated on the hold out data of 3,047 instances.

Advantage of 10-fold over 5-fold

- ✓ Lowers variance in performance estimates.
- ✓ More reliable for small to medium dataset (n=12,188).
- ✓ Standard in public health prediction.

3.10. Development Tools and Computing Environments

In this study, I used different hardware and software tools to develop a binary class child stunting prediction model based on risk factors using an ensemble machine learning techniques. As a hardware tool, I used an Intel Core i7-7200U CPU @ 2.50 GHz or 2.70 GHz, 8 GB of memory, and a 1TB hard disk. The study used software tools such as MS office Word 2010 to write thesis documentation, MS office power point 2010 for thesis presentation, MS office Excel 2010 for understanding the dataset, and Edraw max to draw both conceptual framework and model architecture This study was implemented using Google Colab used to data analysis and model development, Pandas used for data manipulation and numerical computations and NumPy are used to provide faster operation.

CHAPTER FOUR: RESULT AND DISCUSSION

This study enables early stunting detection and prediction, optimizing intervention strategies and Provides data driven policy recommendations for targeted nutrition programs.

4.1. Overview

This chapter presents the experiments conducted to develop the proposed model are discussed here. The experiments were conducted to predict and characterize the contributing factors of child stunting. The results of the developed model are explained and presented, taking into account the training and testing outcomes based on the specified parameter considerations. The experimental evaluation confirms the effectiveness and feasibility of the proposed model architecture. Furthermore, the chapter thoroughly describes the influence of feature selection, SMOTE, hyper parameters optimization and the implementation details of the proposed model and prototype.

4.2. Descriptive Results and Analysis of Distribution Predictive Factors

This section highlights the contribution of selected attributes to target variable based on the feature selection method. These are Child_Age, sex, Exclusive_BF, Maternal_Education, Wealth_Index, Dietary_Diversity, Mother_Occupation, Residence, Region, Place_of_Delivery, Birth_weight, Drought_Exposure, Toilet ,Conflict_Exposure, Mother_Height_cm, Diarrhoea and Displacement.

Stunting Distribution with Sex

Sexes as show below have high significant impact for the contribution rate of children stunting.

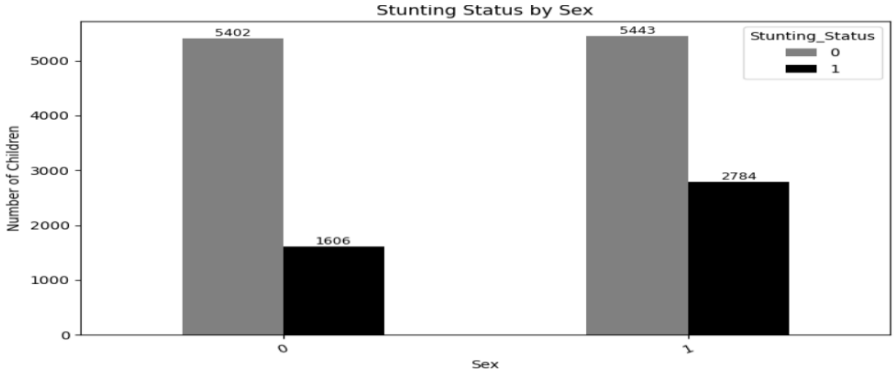


Figure4. 1: Stunting Distribution by Sex

As shown in the above Figure 4.1 male children coded as 1, Out of 8,227 children, 2,784 are stunted and out of 7,008 female children coded as 0, 1,606 children are stunted. On the other hand, the not stunted children, in both groups, were relatively equal as 5,443 not stunted males and 5,402 not stunted females. This indicated that males are about 1.7 times more stunted than females. The Study suggested that male children (54%) were more vulnerable to stunting in early childhood than female (46%). The observed sex based difference is in cycle with the existing nutritional and epidemiological evidence that suggests that biological and behavioral factors may put male children at a higher risk of growth faltering.

Stunting Distribution with Child_Age

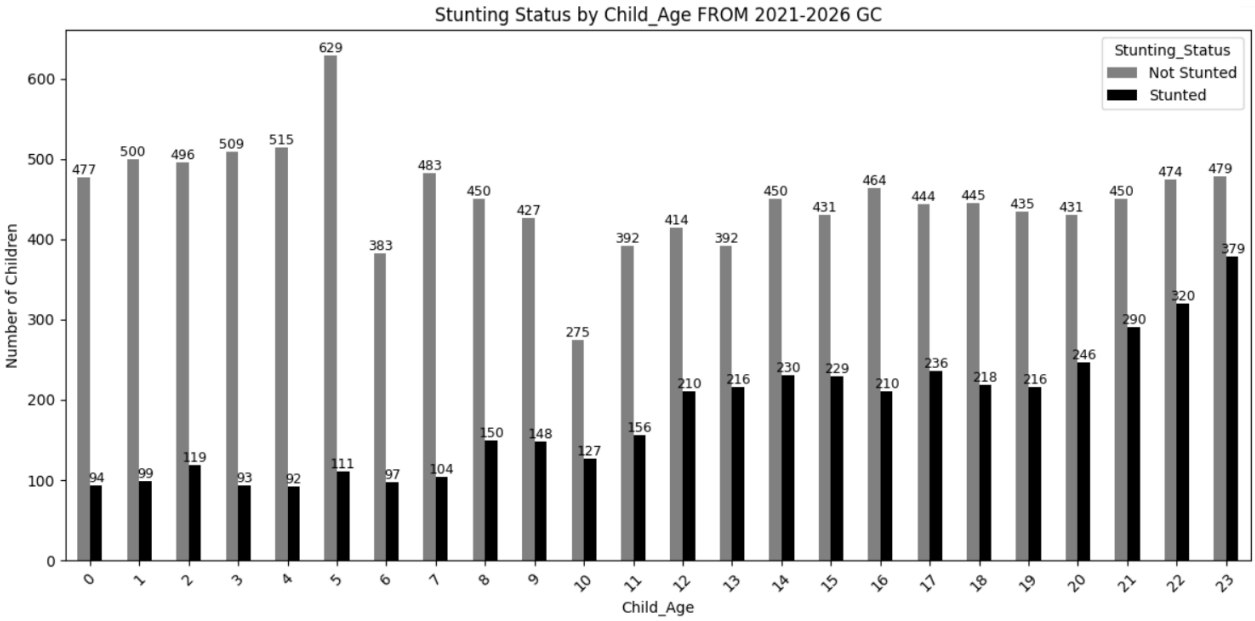


Figure4. 2: Stunting Distributions by Child Age

As shown in the figure 4.2 above when the age increase the probability of stunted also increases in the age categories of 12 to 23 month than less than 12 month. In Age group 0 to11 months out of Total 7,371 children, 1,072 (14.5%) are Stunted and 6,299 (85.5%) are Not Stunted. In the Age group of 12 up to 23 months out of Total 8,327 children 2,394 (28.7%) are Stunted and 5,933 (71.3%) are Not Stunted. Stunting incidence is much more in the 12 up to 23 month children age group (28.7%) compared to the 0 up to 11 month age (14.5%).This show that age groups 12 up to 23 month are critical for stunted expansion, mental and physical growth.

Stunting Distribution with Maternal Education

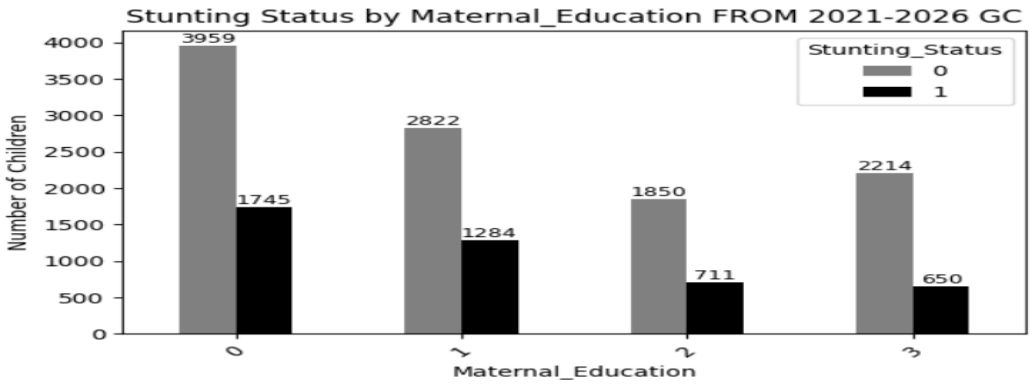


Figure4. 3: Stunting Distributions by Maternal Education

The above Figure 4.3 illustrates the distribution level of maternal education to stunting status. Among children whose mothers had no formal education coded as 0, exactly out of 5,704 children1, 745 (30.6%) were stunted, representing the highest prevalence across all categories. For mothers with primary education coded as 1, the prevalence of stunting declined slightly to 31.3% or 1,284 children out of 4,106. A more pronounced reduction was observed among children of mothers with secondary education as coded 2 out of 2,561, where only 711 (27.8%) children were stunted. The lowest prevalence was recorded among children whose mother's attained higher education coded as 3 out of 2,864, with 650 approximately (22.7%) children classified as stunted. These results indicate that increasing maternal education is related with a reduced chance of child stunting. This indicates maternal education have inverse relationship with stunting status of children. This strategy may reflect enhanced maternal knowledge of nutrition, health seeking behavior, childcare follows, and access to socioeconomic incomes.

Stunting Distribution with Wealth Index

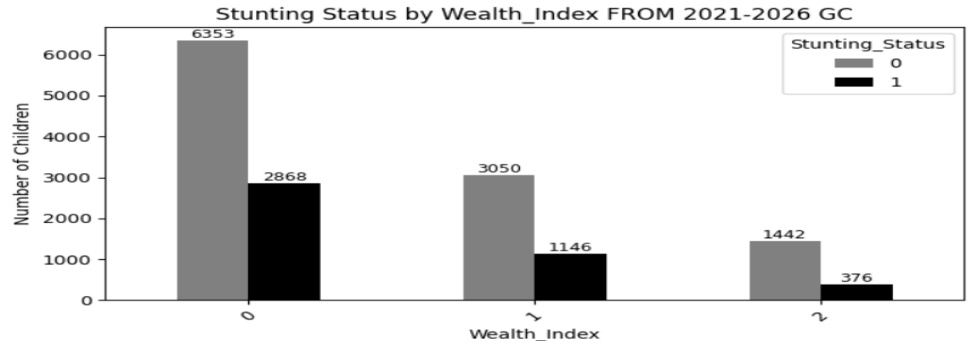


Figure4. 4: Stunting Distribution by Wealth Index

As shown the above Figure4.4, stunting status across wealth classes tells a clear socioeconomic level in child nutritional outcomes. Children from poor households coded as 0 indicates the highest incidence of stunting, out of 9,221 Children, 2,868 31% were stunted cases compared to 6,353 non-stunted. In the other hand, middle income households coded as 1 show a reduced scale, out of 4,196 children 1,146 27.3% were stunted and 3,050 not stunted. Children from rich families member coded as 2 indicates lowest incidence, out of total 1,818 children only 376 20.7% children were stunted and 1,442 were not stunted. This study shows that as household wealth increases, the probability of stunting incidence decreases, highlighting the protective role of economic resources in child growth and development. The results highlight the significance of addressing poverty related causes of underfeeding, as wealth disparities remain a significant forecaster of stunting among under-two children.

Stunting Distribution with Residence

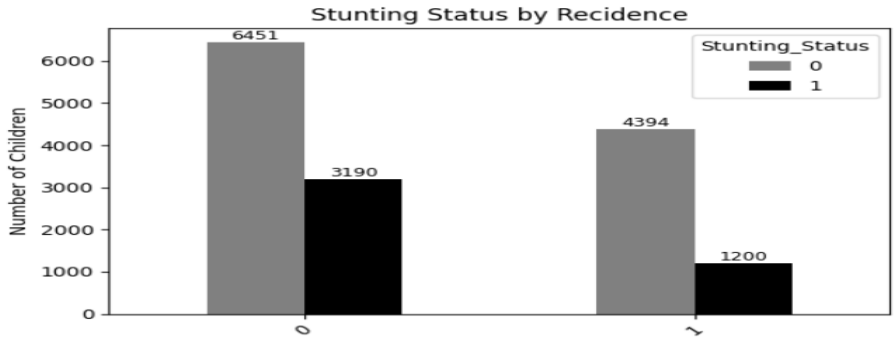


Figure4. 5: Stunting Distribution by Residence

As shown in figure 4.5 out of 9,641 the children who love in rural residence 3,190 (33.08%) were stunted and 6451 are not stunted.in the other hand, out of 5,594 the children who love in urban residence 1,200 (21.45%) were stunted and 4,394 are not stunted. The probability of stunted rate is increased in rural residence than in urban residence due to different factors. These studies indicate that more intervention is required to apply in rural areas to prevent children from stunting with expanding multi sectorial activities with in different sectors.

Stunting Distribution with Region

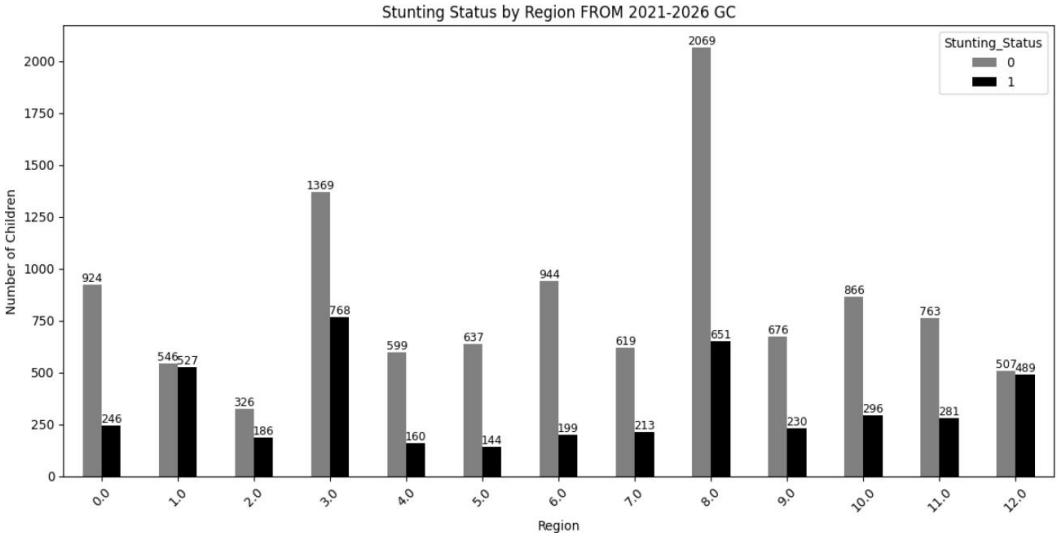


Figure4. 6: Stunting Distribution by Region

As shown in fig 4.6 Afar Region which is indicated the highest number of stunted rate, and is as a result the biggest hotspot for child stunting, out of 1,037 children 546 50.8% are stunted and 527 are not stunted. This implies major problems with nutrition, conflict exposure, drought exposure, food security and maternal-child health. Tigray also have the 2rd rank have highest stunted rate next to Afar, out of 996 children 489 49% are stunted and 507 are not stunted and Sidama also have the 3rd rank have highest stunted rate next to Tigray, out of 511 children, 186 36.3% are stunted and 326 are not stunted. Amhara coded as 3 have the 4th highest overall number of children out of 2,137 children, 768 35.9% are stunted and 1,369 are not stunted, suggesting interventions must be scaled up for a large population. Regions such as Oromia, South West, Somali, Hareri, SNNPR and Benishangul-Gumuz have moderate total amounts, but proportionally more risks, given their drought, pastoral economies and poor infrastructure. Urban regions such as Gambela, Dire Dawa and Addis Ababa have low amount, but stunting perseveres due to socioeconomic differences, poverty, and lack of access to reasonable nutritious food. Regions affected by conflict like Amhara and Tigray have high records but the causes are conflict and service interferences.

Stunting Distribution with Birth weight

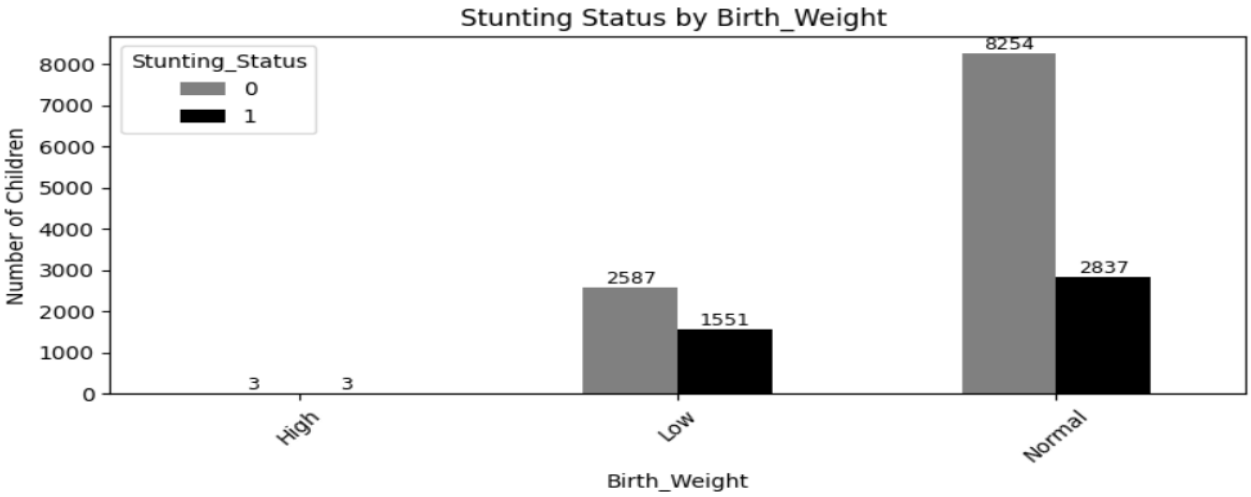


Figure4. 7: Stunting Distribution by Birth Weight

The Figure 4.7 indicated that probability of birth weight is strongly associated with child stunting. Among children who have low birth weight, have 37.5% of highest Stunting incidence, compared to 25.6% among those with normal birth weight. While the high birth weight group was very small, half of the children in this category were stunted. These findings underline birth weight as a serious biological factor of child growth outcomes, reflecting maternal health condition and prenatal nutrition.

Stunting Distribution with Dietary Diversity

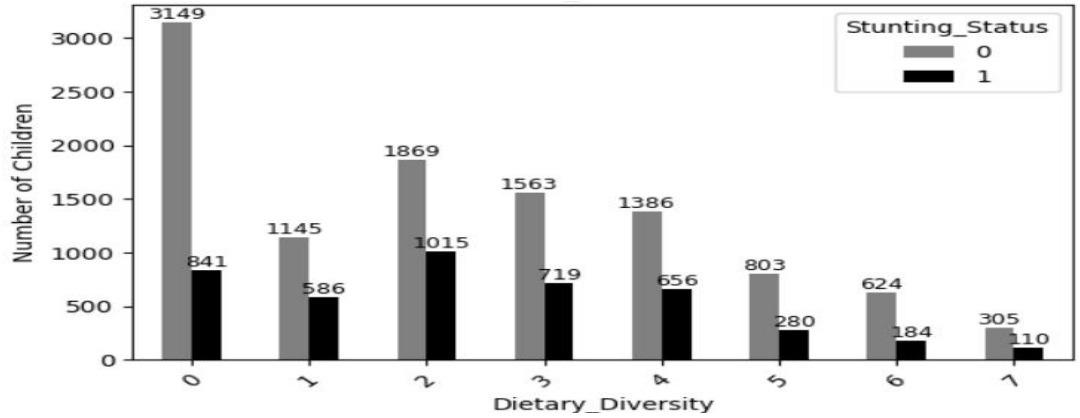


Figure4. 8: Stunting Distribution by Dietary Diversity

The Figure4.8 indicates that probability of stunting is the highest for children with low dietary diversity of 0 up to 3 food groups, out of 10,887 children more than 3,161(29%) children are stunted. In the medium group of 4 up to 5 food groups, the prevalence of stunting is the same,

30% However, with a lesser absolute number of stunted children 936 out of total 3,125 children. Among children who are consuming a high dietary diversity 6 up to 7 food groups, stunting prevalence decreases to 24% and 294 are stunted out of total 1,223 children, suggesting a protective effect of dietary diversity. On a national level, the data indicate that the higher the number of food groups consumed, the lower the rate of stunting, relative to the population and in absolute terms.

Stunting Distribution with Drought Exposure

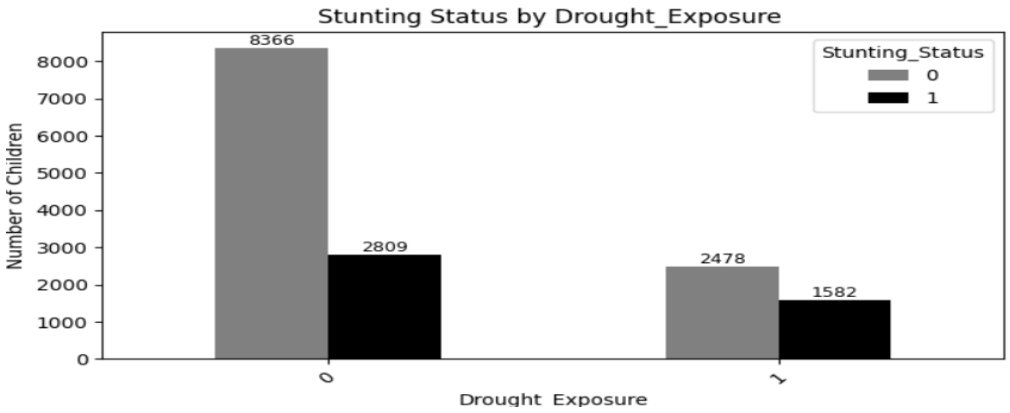


Figure4. 9: Stunting Distribution by Drought Exposure

The above Figure4.9 shows prevalence of stunting is relatively low 25% for children not exposed to droughts than it is for those exposed. In children who were exposed to drought, the proportion of stunted children increases noticeably to 39%, indicating that drought conditions affect the nutrition situation of children negatively. The number of case among non-exposed is higher 2,809 case out of 11,175 children, but this is because the population is larger. The comparative risk was more in the group exposed to droughts out of 4,060 children 1,582 were stunted. This indicates that drought exposure was a significant determinant of stunting, affecting on dietary sufficiency and resilience of households.

Stunting Distribution with Conflict Exposure

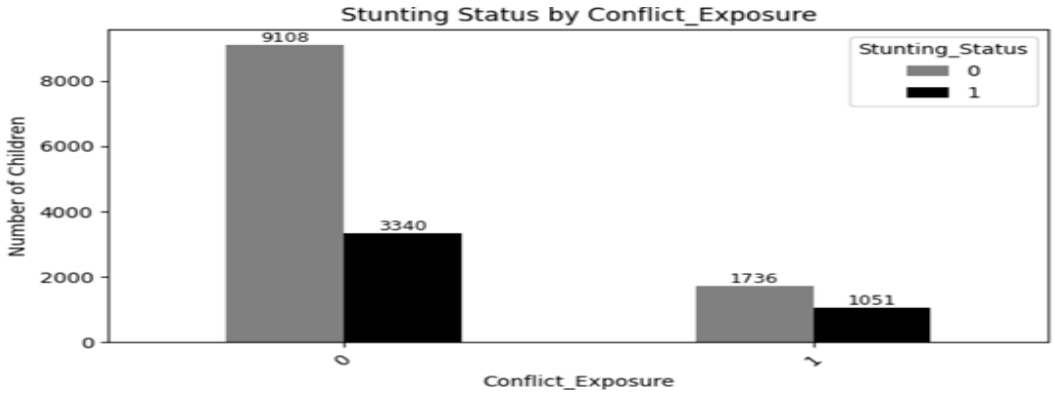


Figure4. 10: Stunting Distribution by Conflict Exposure

Children exposed to conflict coded as 1 had higher chance to stunting compared to those not exposed to conflict coded as 0. Whereas the total number of children is lower in conflict-affected areas due to displacement, the relative weight of incidence in stunting increases**.** Specifically, in non-conflict exposed members, out of total 12,448 children 3,340(26.8%) were stunted, whereas in conflict exposure houses, out of total 2,787 children 1,051(37.7%) were stunted. This demonstrates that conflict exposure exacerbates stunting incidence, highlighting the susceptibility of children in insecure environments.

Stunting Distribution with Diarrhoea

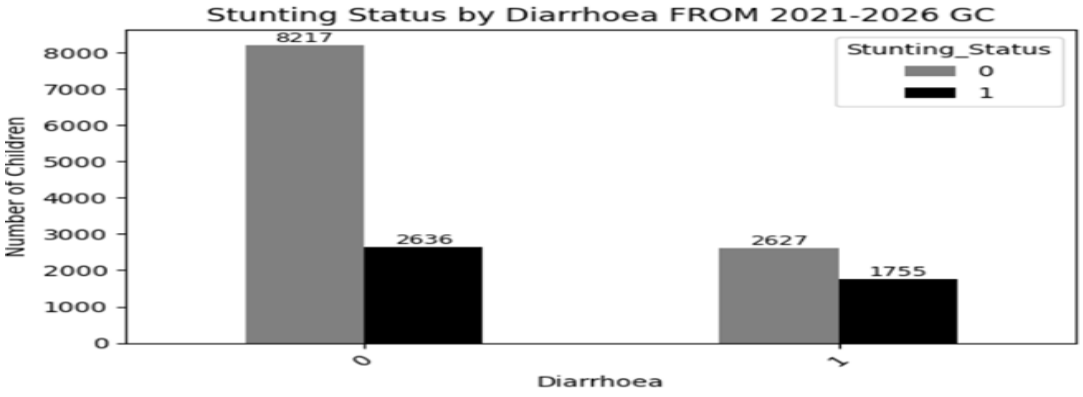


Figure4. 11: Stunting Distribution by Diarrhoea

Children who did have Diarrhoea coded as 1 indicates a higher incidence of stunting compared to those who did not had. Among children did not have diarrhea, out of total 10,853 children 2,636 (24.3%) were stunted. In other hand, among children with Diarrhoea, out of total 4,382

instance 1,755 (40.0%) were stunted. This insight highlights infectious disease aggravating stunting and emphasizes the importance of preventive health interventions.

Stunting Distribution with Exclusive Breast Feeding

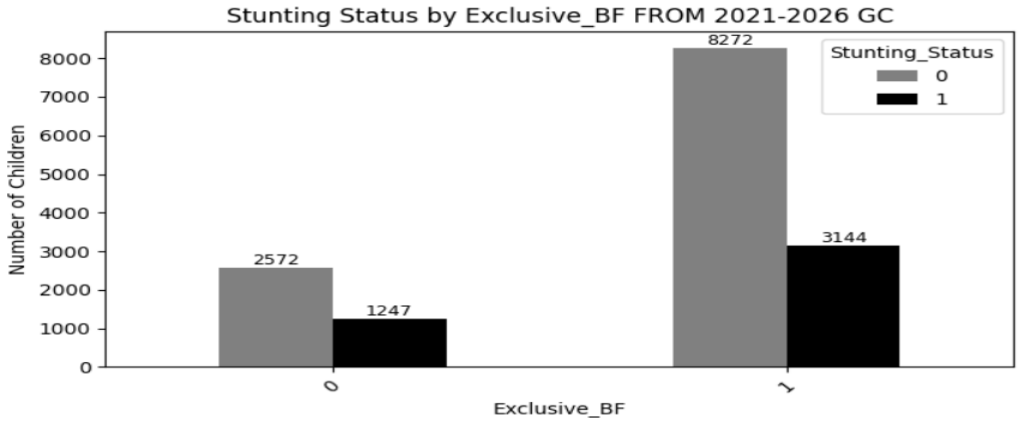


Figure4. 12: Stunting Distribution by Exclusive Breastfeeding

The above Figure4.12 show the contribution Exclusive breastfeed for Children who were only breastfeed coded as 1 had a lower incidence of stunting than those who were not breastfeed. Of children who were not exclusively breastfed, 1,247 (32.7%) were stunted and 3,819 children were not exclusively breastfed. In the other hand, 11,416 children were exclusively breast fed with 3,144(27.5%) stunted children. This discovery highlights the need to exclusively breastfeed as an essential life stage intervention to reduce undernutrition and ensure healthy growth.

Stunting Distribution with Toilet

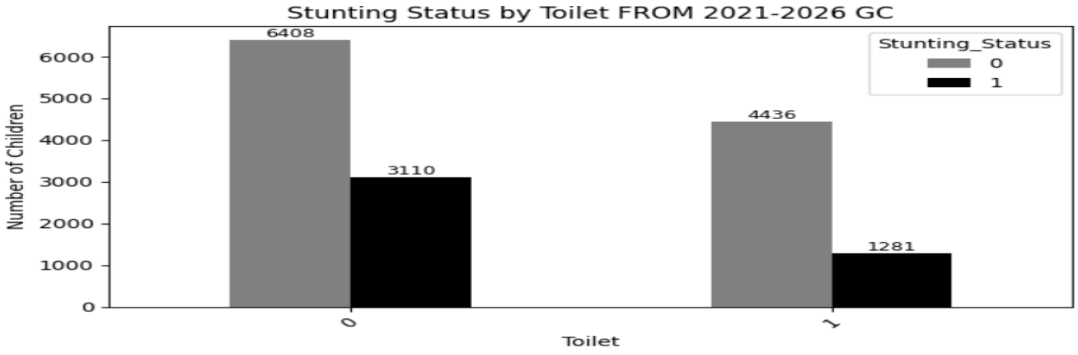


Figure4. 13: Stunting Distribution by Toilet

Higher prevalence of stunting is observed among children from households without toilet facility (coded as 0) compared with children who have access to toilets (coded as 1). In particular, 3,110

(32.7%) out of 9,518 children in the households without toilets were stunted. Among those households that had toilets, 1,281 (or 22.4%) children were stunted out of 5,717 children. This suggests that the presence of sanitation facilities is also associated with a lower proportion of stunted children, which may be related to lower incidence of infections and better hygiene and to lower burden of Diarrhoea disease.

Stunting Distribution with Mother Occupation

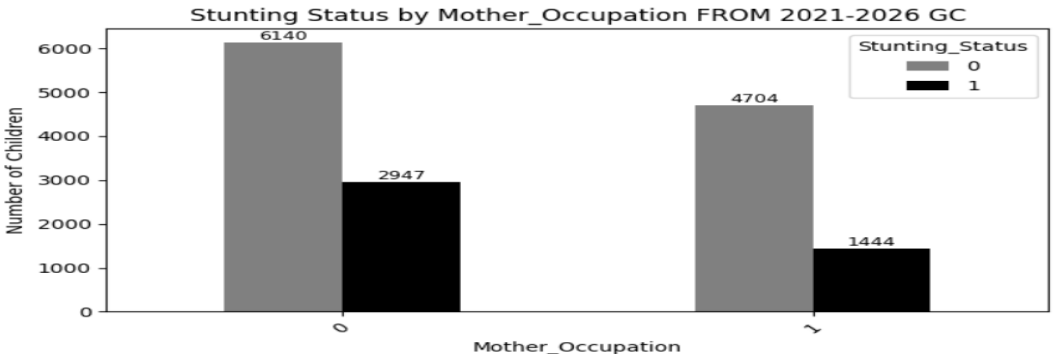


Figure4. 14: Stunting Distribution by Mother Occupation

Children of mothers who have no occupation coded as 0 have a higher proportion of stunted children than do children of mothers who are employed, coded as 1. In particular, 2,947 (32.4%) of the 9,087 children were stunted among non-working mothers. Among working mothers 6,148 stunted children were identified out of which 1,444 (23.5%) were identified as stunted children. This indicates that employment of mothers may have indirect effects (improved income, food security, health care etc.) that also reduce the risk of stunting.

Stunting Distribution with Displacement

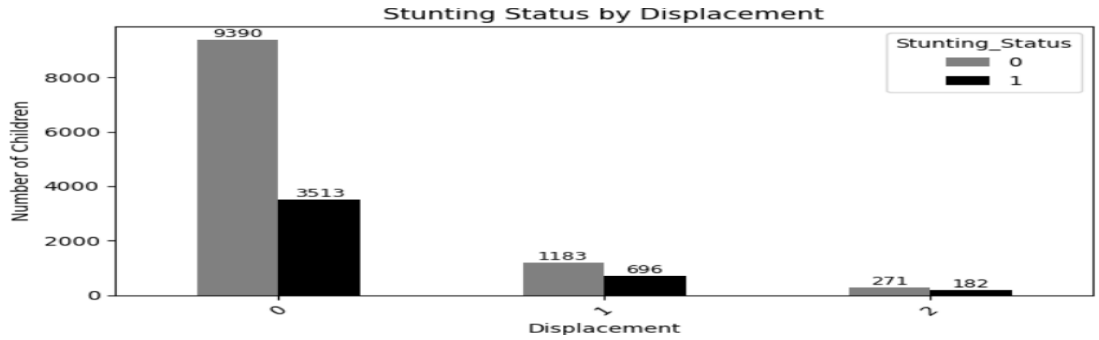


Figure4. 15: Stunting Distribution by Displacement

As seen in the Figure4.15 displacement and child stunting are highly correlated. There was higher rate of stunting among Internal and External displaced households as compared to non-displaced households, from 27.2% to 37.0% and 40.2% respectively, regionally. This demonstrates the displacement as a structural determinant of child nutrition, and the cumulative vulnerability of child populations in crisis.

Over All Distribution of Stunting Status

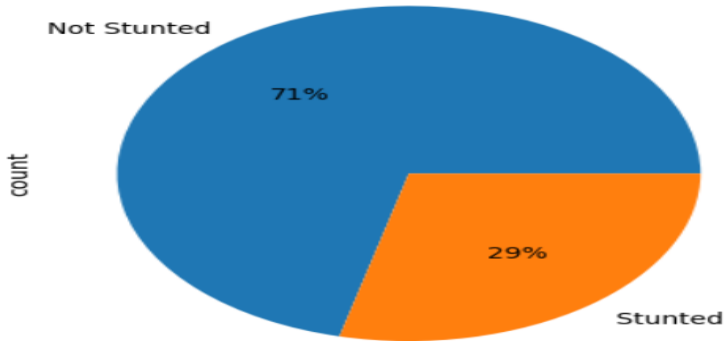


Figure4. 16: Stunting Status Distribution

As shown the above Figure 4.16 nearly 1/3 (29%) of children under two years of age study population ware stunted, while the majority 71% were not stunted. This indicates stunting remains a significant health concern, affected almost one third (one in three) of Ethiopian children from 2021 up to 2026. The incidence highlights the magnitude of the problem and provides a baseline for further risk factor analysis and predictive model. This study indicates that to end stunting there must be needed an intervention with much more sectors and policy reforms. The overall insight of this study in Ethiopia is shaped by biological factors (child age, sex, birth weight and Diarrhoea), socioeconomic factors (maternal education, occupation, residence) and structural exposures such as region drought, conflict, displacement and sanitation. The findings highlight that multi sectorial intervention such as nutrition, maternal education, poverty reduction, health service and resilience to shocks are essential to reduce stunting.

4.3. Implementation Details

This section highlights how the experiment is conducted, the first task done after data preprocessing and class balancing is significant feature selection to identify the best prediction model and model explainability. To construct a prediction model for stunting based on risk

factors the study used important features and the best appropriate ensemble machine learning algorithm. In this research, it conducted five experiments using ensemble machine learning algorithms with relevant features selected by using Boruta feature selection method.

4.3.1. Handling missing values

```
df.isnull().sum()
```

	0
Child_Age	0
Sex	0
Birth_Weight	1
Diarrhoea	0
Exclusive_BF	0
Maternal_Age	0
Maternal_Education	0
Maternal_Anaemia	1
Mother_Occupation	0
Marital_Status	0
Mother_Height_cm	1
Mother_BMI	0
Antenatal_Visits	0
Residence	0
Wealth_Index	0
Dietary_Diversity	0
Water	4
Toilet	0
Place_of_Delivery	1
Region	1
Market_Access	0
Drought_Exposure	0
Conflict_Exposure	0
Displacement	0
Food_Aid	0
HAZ	1
Stunting_Status	0

dtype: int64

```
df.isnull().sum()
```

	0
Child_Age	0
Sex	0
Birth_Weight	0
Diarrhoea	0
Exclusive_BF	0
Maternal_Age	0
Maternal_Education	0
Maternal_Anaemia	0
Mother_Occupation	0
Marital_Status	0
Mother_Height_cm	0
Mother_BMI	0
Antenatal_Visits	0
Residence	0
Wealth_Index	0
Dietary_Diversity	0
Water	0
Toilet	0
Place_of_Delivery	0
Region	0
Market_Access	0
Drought_Exposure	0
Conflict_Exposure	0
Displacement	0
Food_Aid	0
Stunting_Status	0

dtype: int64

Figure4. 17: Missing values before and after Imputation

To address missing data of continuous and numerical features such as Birth Weight**,** Mother Height**,** and HAZ, appropriate imputation techniques were applied. Missing value was handled using mean imputation for numeric value features, whereby the average of the detected values was replaced for the missing entries. This technique conserves the overall distribution of the data and decreases bias in subsequent arithmetic investigates. In other hands for categorical variables like Maternal-Anaemia**,** Water**,** and Region, missing values were imputed using the mode, which replaces missing records with the most frequently occurring class. This system confirms that the categorical structure of the data is maintained while decrease the risk of distorting class distributions. By employing mean imputation for numerical variables and mode imputation for

categorical variables, the dataset was reduced complete and suitable for robust analysis in the stunting prediction models.

4.3.2. Correlation Matrix

Correlation matrix used to show the relationship among independent feature and with stunting status of under-two child stunting.

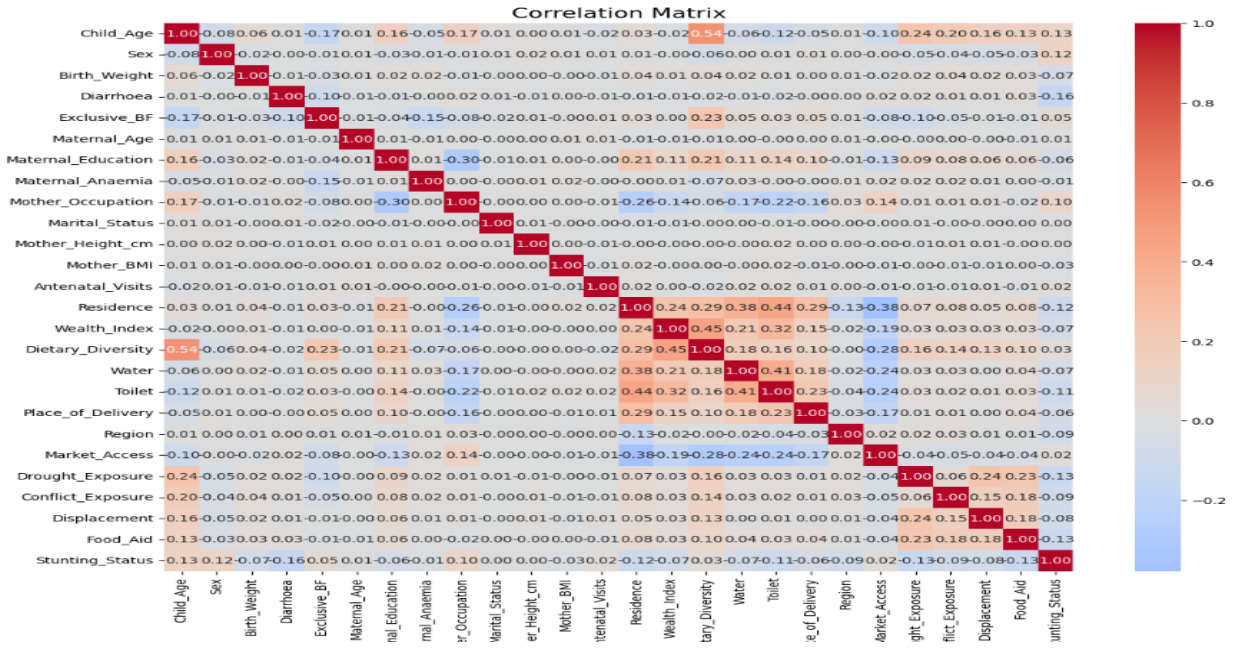


Figure4. 18: Correlation matrixes among attributes and stunting status

As the above fig 4.8 indicate that, the correlation matrix shows that the relationship between independent features with each other and with dependent feature. The highest value of the feature is more important and less value of features which have less important for the target class. From the correlation matrix, the relationship between Child Age and Stunting Status was the strongest among all predictors, and have Positive correlation indicates as child age increases, the likelihood of being stunted also increases. The correlation between Birth Weight and stunting Status was one of the second strongest forecasters. Birth Weight had Negative association; shown lower birth weight was related with higher stunting risk. The correlation between Dietary Diversity and stunting Status have moderate but important; lower dietary diversity was associated with higher stunting risk. Children consuming fewer food groups were more likely to be stunted.

The correlation between Maternal Education and Maternal Height with stunting Status was negative; shows higher maternal education and Maternal Height were related with lower stunting risk. Maternal height reflects long term nutritional status and intergenerational health. Shorter mothers often experienced undernutrition in infant themselves, which biologically influences their children to growth undecided.

4.3.3. Dataset Splitting

The train-test split procedure is used to estimate the performance of machine learning algorithms when they are used to make predictions on data not used to train the model. In this study, first, I split the 15,235 total instances of the dataset into a train, and test size split proportion, such as a 70%, 30% to 80%, 20% test split size proportion, with the help of a ratio-based random splitting technique. Because of the lower performance in the train test split with a proportion of 70%, 30%, I used only a proportion of 80%, 20% for the train test split to develop the proposed model. In 80%, 20% proportion, out of 15,235 instances, 12,188 (80%) instances are training datasets that are used for building a prediction model, and 3,048 (20%) instances are test datasets that are used to test the performance of the predictive model of the stunting .

4.3.4. SMOTE

The Synthetic Minority Over sampling Technique (SMOTE) was used to overcome class imbalance problem in the training data set. After apply SMOTE total dataset is improved from 15,235 to 17,380.Before balancing, the dataset appears to have a disproportionate number of samples in each of the two outcome classes, which may result in biased modeling and less predictability. After applying SMOTE, both classes coded as one for non-stunted and one for stunted were resampled to make equal amount of 8,690 rows each. This balancing confirms that the machine learning models are trained on a dataset with proportional representation, in this manner minimizing bias toward the majority class and improving the robustness of stunting prediction. The balanced dataset provides a more reliable basis for ensemble machine learning approaches, specifically in contexts where malnutrition outcomes are inherently skewed in population distributions.

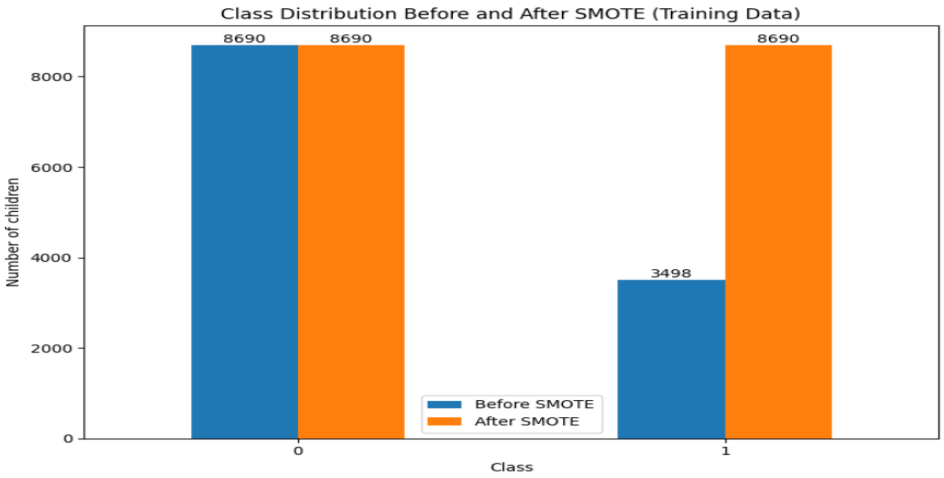


Figure4. 19: class distribution before and after SMOTE

4.3.5. Hyperparameter Optimization

The performance of the ensemble machine learning algorithm highly depends on the selection of parameters, which has always been an essential step .Optuna is a method of HPO that constructs and asses the model scientifically for each for arrangement of parameters used in the model. Here, the study used Optuna and stratified cross validation of 5 folds to select tuning parameters for an ensemble machine learning algorithm and to prepare distribution of training and testing data. The tuning parameters for each ensemble machine learning algorithm are identified by Optuna for all experiments as seen Table 4.1below.

Table4. 1: Hyperparameter Tuning

No	Models/Algorithm	Parameters	Values
1	Random forest	n_estimators	100-500/499
		max_depth	3-20/16
2	Gradient boosting	n_estimators	100-400/333
		max_depth	3-6/4
		learning_rate	0.01-0.2/0.06
3	Extreme gradient boosting	n_estimators	100-500/448/356
		max_depth	3-8/4
		learning_rate	0.01-0.2/0.077
		Subsample	0.6-1/0.83
		colsample_bytree	0.6-1.0/0.84

		eval_metric	Logloss
4	Light boost	n_estimators	100-500/466
		num_leaves	20-100/20/39
		learning_rate	0.01-0.2/0.01
5	Catboost	Iterations	200-500/389
		learning_rate	0.01-0.2/0.13
		Depth	4-8/4

4.4. Model Comparison

To build a predictive model for stunting prediction based on risk factors, the study carried out a total of five experiments using the features that were chosen as input. It performed five experiments, experiment #1 to experiment #5, using the ensemble machine learning models such as Random forest, Gradient boosting, Extreme gradient boosting, CatBoost and Light Boost to create a prediction model. As a result, I compared the performance of algorithms to predict the Stunting status using ensemble machine learning model. The study used overall accuracy and the AUC-ROC curve as an evaluation for predictive model comparison. According to the overall accuracy, the prediction algorithm that recorded the highest performance is used as the best model for predicting of Stunted and not- Stunted.

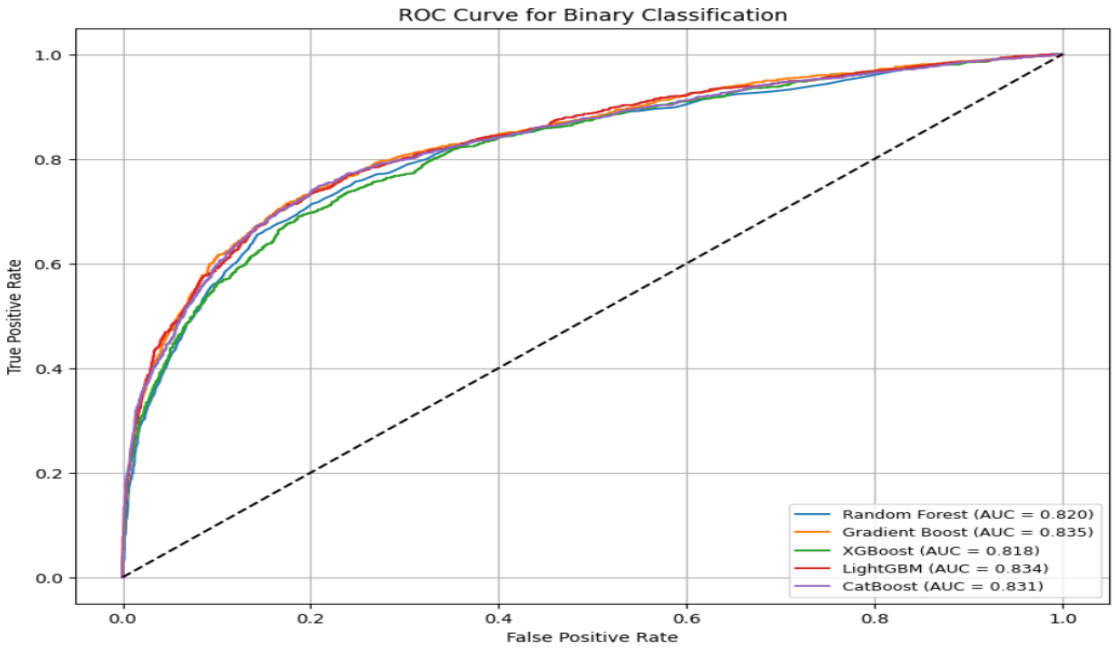


Figure4. 20: AUC- ROC curve after HPO

As the figure 4.23 ROC curves indicate that all the models evaluated have relatively high true-positive rates (TPR) for a wide range of classification thresholds while maintaining moderately low false positive rates (FPR). Gradient Boosting (AUC = 0.835), Light Boost (AUC = 0.834) and CatBoost (AUC = 0.831) had high AUC values indicating that these models had better performance in correctly discriminating the stunted children (True positives) while not misclassifying the non-stunted children (False positives) significantly. In specific, the ROC curves growth sharply in the low FPR area suggesting that the large portion of stunted cases can be captured with a fairly low FPR. The performance was high although the TPR was slightly lowered for models with slightly lower AUC, such as Random Forest and XGBoost, for the same FPR. Overall, findings indicate the balance between sensitivity and specificity was better, particularly in a public health setting with the goal of targeting children who may benefit from intervention services, but not over-intervention on children who were not at risk.

Table4. 2: The Overall Model Performance of Experiment

Evaluation metrics	Random Forest	Gradient Boost	XGBoost	Light Boost	Cat Boost
Before SMOTE and HPO					
Accuracy	79.75	81.42	79.45	80.47	80.89
Recall	46.80	49.38	50.50	49.94	50.72
Precision	74.64	79.45	71.02	75.08	76.13
F-score	57.53	60.91	59.03	59.98	60.88
ROC Curve	80.74	82.66	81.02	82.41	82.82
Rank	4	1	5	3	2
After HPO					
Accuracy	80.95	81.72	81.76	81.49	81.72
Recall	80.95	81.72	81.76	81.49	81.72
Precision	80.21	81.09	81.15	80.81	81.10
F-score	79.92	80.81	80.79	80.66	80.76
Auc curve	81.96	82.89	82.81	82.87	82.96
Rank	5	3	1	4	2

This Table show that the model training performance of two experiments. Those are model training before SMOTE and HPO and model training after hyper parameter optimization (HPO). In both case the research evaluates all the models using evaluation matrix. As compare the two

experiments model training before hyper parameter tuning and model training after hyper parameter tuning the performance of the models are difference values. The performance the models before hyper parameter tuning are smaller than that of after hyper parameter optimization. This result indicates that hyper parameter optimization is essential to enhance the models performance. So the research is conducted by using after HPO. Table 4.2 shows that the XGBoost had the overall best-predicted performance, with an accuracy of 81.76%. While the Random forest model had the relatively less performance, with an accuracy of 80.95%. XGBoost registered the highest accuracy of 81.76% and 82.81% for the ROC curve, and the 2nd mode having the highest accuracy is Cat Boost with an accuracy of 81.72% and a ROC curve of 82.96%. Gradient Boost is the 3rd mode with accuracy of 81.72% and 82.89% ROC curve. Light Boost mode is the 4th with the accuracy of 81.49% and 82.87% ROC curve. Random forest is the last with accuracy of 80.95% and 81.96% ROC curve. So, XGBoost is selected as the best ensemble machine learning as compared to another ensemble machine learning according to the results of Table 4.2.

4.5. Test Performance of Models Result and Analysis

In this phase, the study applied and tested Random Forest, Gradient Boosting, XGBoosting, Cat Boost trees and Light Boost to predict Stunting. To build a predictive model, 15,235 instances and 17 features with 80 by 20 ratios were used. To evaluate the performance of the model, 5-fold cross validation is used due to its relatively low bias and variations. This means that the data is randomly divided into five equal parts. That is four folds for training and one fold for validation.

Experiment# 1: Random Forest

In this experiment, the Random Forest was used 17 best important feature as input chosen from the above experiment to develop a predictive model and achieved an accuracy of 80.95%, precision of 80.21%, recall of 80.95%, F1-score of 79.92% and AUC of 81.96% .

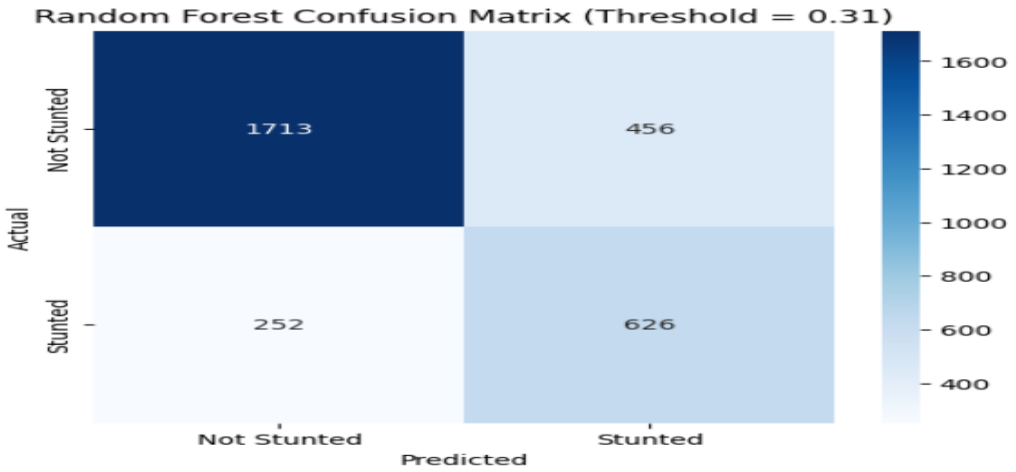


Figure4. 21: Confusion Matrix of Random Forest

Table4. 3: Confusion Matrix of Random Forest

Actual class	Predicted Not Stunted	Predicted Stunted	Total
Not Stunted	1713 (True Negative)	456 (False Positive)	2169
Stunted	252 (False Negative)	626 (True Positive)	878

Based on Table 4.3, the Random forest model was tested on 3,047 samples. Among these, 2,339 instances were correctly classified or predicts, while 708 instances (were misclassified. Specifically, among 878 children who were actually stunted, 626 were correctly identified (true positives) and 252 were incorrectly classified as not stunted (false negatives). Among 2169 children who were actually not stunted, 1,713 were correctly identified (true negatives) and 456 were incorrectly classified as stunted (false positives). These results indicate that the Random forest model achieved relatively low performance, with better accuracy in identifying non-stunted children, though some stunted cases were missed. All correctly predict instances have with an accuracy of 76.8%, and the wrong prediction had an accuracy of 23.2%.

Experiment# 2: Gradient boosting similarly, the study has also built a predictive model using a gradient boosting classifier, by using 17 relevant features. As illustrated in figure 4.22 below, this algorithm has a performance of ROC 82%, 81.72% of accuracy, precision of 81.09%, recall of 81.72% and F1-score of 80.81%.

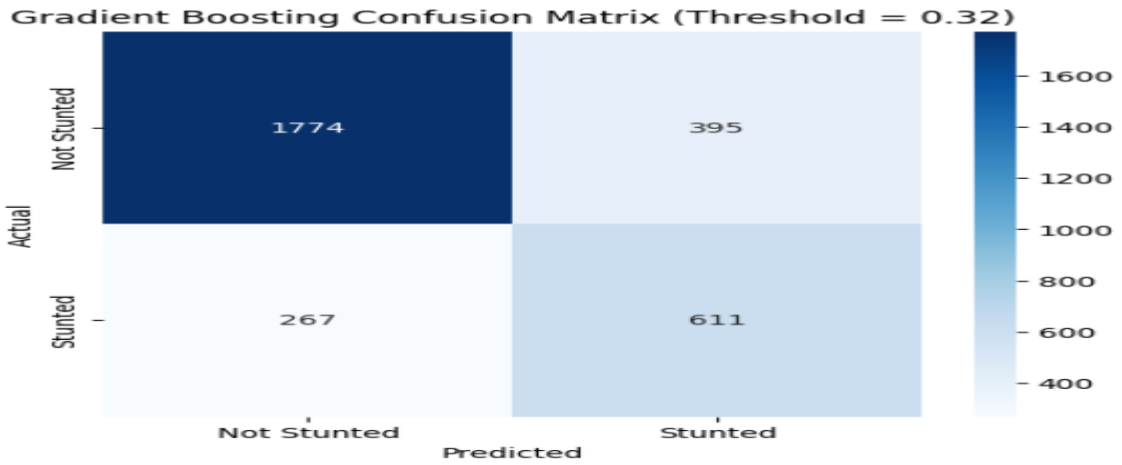


Figure4. 22: Confusion Matrix of Gradient Boosting

Table4. 4: Confusion Matrix of Gradient Boosting

Actual class	Predicted Not Stunted	Predicted Stunted	Total
Not Stunted	1,774(True Negative)	395(False Positive)	2,169
Stunted	267(False Negative)	611(True Positive)	878

Based on Table 4.4 above, in gradient Boost there are 3,047 samples of data were used for test the model. Among the test sample data 2,385 were correctly classified and the remaining 662 samples of data were incorrectly classified. Out of 878 sample 611 instances are correctly predicts as having stunted (True Positive), whereas 267 instances is incorrectly predicts (False Negative) as not having stunted. 1,774 instances are correctly predicts as not stunted (True Negative), whereas 395 instances are incorrectly predicts as stunted (False Positive) stunted out of total 2,169 case. The correctly predicted instances have an accuracy of 78.3% and the wrongly predicted value was 21.7%.

Experiment# 3: Extreme gradient boosting similarly, the study also built a predictive model using an extreme gradient boosting (XGBoost) classifier, by using 17 best relevant features. As illustrated in Figure 4.14 below, this algorithm has a performance 82.81% of ROC, 81.76% accuracy, precision of 81.15%, recall of 81.76% and F1-score of 80.79%.

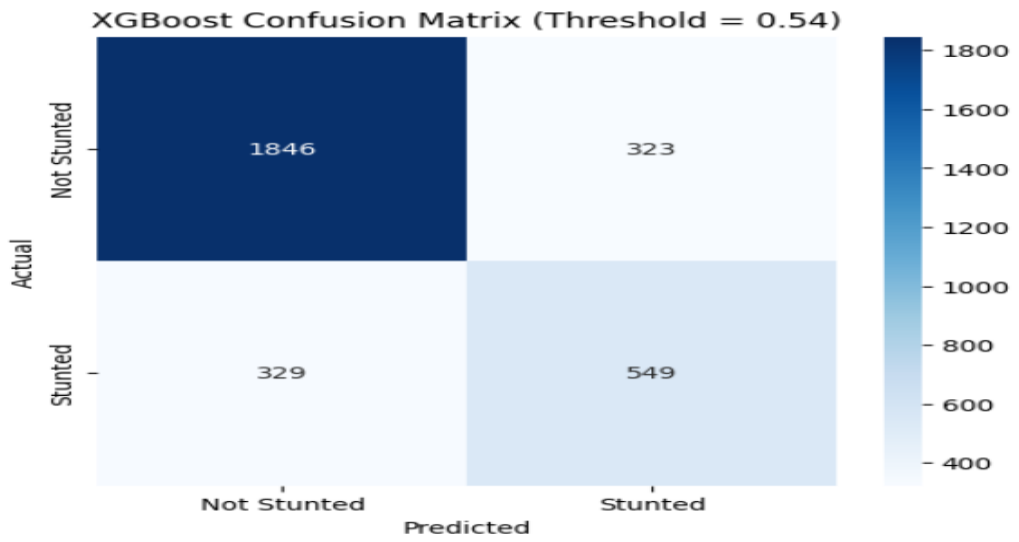


Figure4. 23: Confusion Matrix of Extreme Gradient Boosting

Table4. 5: Confusion Matrix of Extreme Gradient Boosting

Actual class	Predicted Not Stunted	Predicted Stunted	Total
Not Stunted	1846(True Negative)	323(False Positive)	2169
Stunted	329(False Negative)	549(True Positive)	878

Based on Table 4.5, the XGBoosting model was tested on 3,047 instances. Out of these, 2395 instances were correctly classified, while 652 instances were incorrectly predicted. Specifically, out of 878 children who were actually stunted, 549 were correctly identified (True Positives) and 329 were incorrectly predicted as not stunted (False Negatives). Among 2169 children who were actually not stunted, 1,846 were correctly identified (True Negatives) and 323 were incorrectly classified as stunted (False Positives). The results show that the XGBoosting model successfully predicted some not stunted and some stunted children were not predicted, but it is relatively good for predicting children who are not stunted. The correctly predict instances an accuracy of 78.6%, and the wrong predict have an accuracy of 21.4%.

Experiment# 4: Light Boost Similarly, the study have built a predictive model using a Light Boost classifier by using relevant features that are selected for building a model by using Boruta feature selection method. In this experiment, as you can see figure 4.15 below, this model has the best performance of 82.87% ROC, accuracy of 81.49%, precision of 80.81%,recall of 81.49% and F1-score of 80.66.

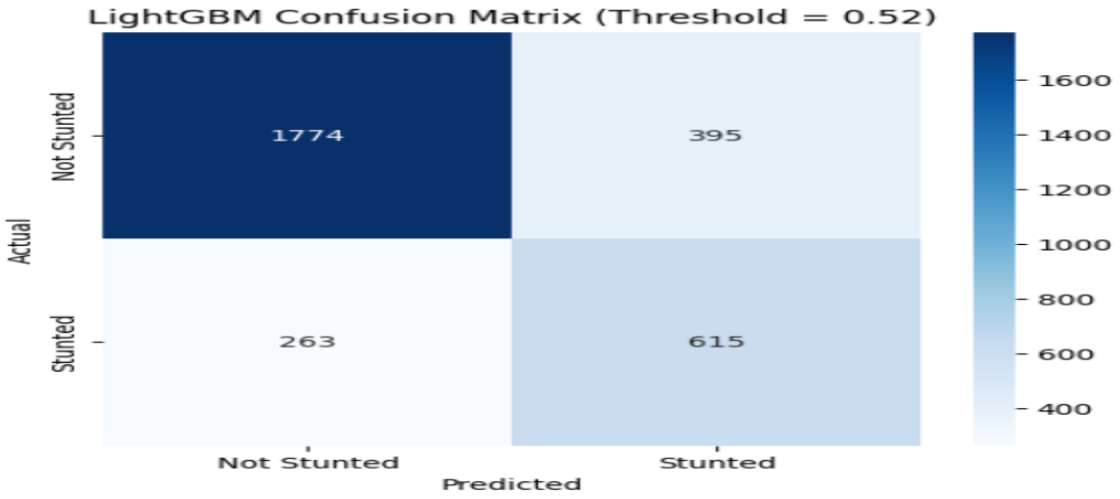


Figure4. 24: Confusion Matrix of Light Boost

Table4. 6: Confusion Matrix of Light Boosting

Actual class	Predicted Not Stunted	Predicted Stunted	Total
Not Stunted	1774(True Negative)	395 (False Positive)	2169
Stunted	263(False Negative)	615 (True Positive)	878

Based on Table 4.6, the light Boosting model was tested on 3,047 instances. Among these, 2389 instances were correctly classified, while 656 instances were misclassified. Explicitly, among 878 children who were actually stunted, 615 were correctly identified (True Positives) and 263 were incorrectly classified as not stunted (False Negatives). Among 2,167 children who were actually not stunted, 1,774 were correctly identified (True Negatives) and 393 were incorrectly classified as stunted (False Positives). These results indicate that the light Boost model achieved relatively strong performance, with well accuracy in identifying non-stunted children; however, some stunted children were missed with predicted accuracy of 21.55% and correct predicted of instances have an accuracy of 78.45%.

Experiment# 5: Cat Boost Similarly, the study built a predictive model using a Cat Boost classifier by using relevant features that are selected for building a model by using sequential step-backward feature selection. In this experiment, as seen in figure 4.15 below, this model has the best performance of 81.96% ROC, accuracy of 80.95%, precision of 80.2% recall of 80.95%, and F1-score of 79.92%.

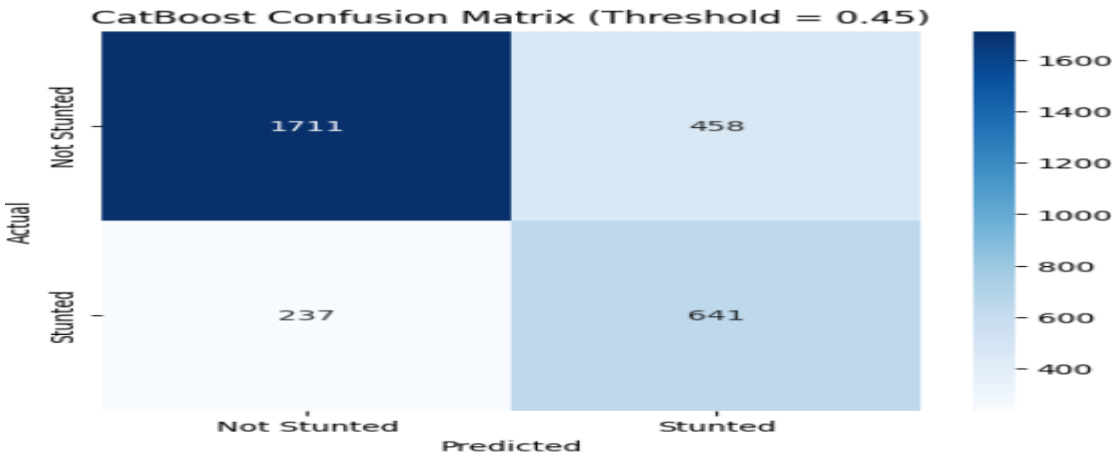


Figure4. 25: Confusion Matrix of Cat Boost

Table4. 7: Confusion Matrix of Cat Boosting

Actual class	Predicted Not Stunted	Predicted Stunted	Total
Not Stunted	1711(True Negative)	458(False Positive)	2169
Stunted	237(False Negative)	641(True Positive)	878

Based on above Table 4.7, the Cat Boost model was tested on 3,047 instances. Of these, 2352 were classified correctly and 695 were misclassified. Obviously, from 878 children who were actually stunted, 641 were correctly identified (True Positives) and 237 were incorrectly classified as not stunted (False Negatives). Of the 2169 children who were actually not stunted, 1,711 children were correctly identified (True Negatives) and 458 children were incorrectly identified as stunted (False Positives). These results indicate that the Cat Boost model achieved relatively strong performance, with well accuracy in identifying non-stunted children; but, some stunted issue were missed. The overall correctly predicted instances had an accuracy of 77.2%, and the wrongly predict accuracy of instances was 22.8%.

4.6. Feature importance

In this study, the most determinant factors were identified using the best-performing algorithms though their feature importance scores (see Table 4.8). Among the above five experiments, XGBoost achieve the best performs across evaluation metrics. To identify the most important predictors of stunting, random forest feature importance techniques was applied. The feature with the highest value represent the most predicted factor, while with the lowest value is the least weak factor.

Table4. 8: Feature Importance Ranking

No	Feature	Values	No	Feature	Values
1.	Child_Age	0.130	10.	Diarrhoea	0.031
2.	Birth_weight	0.113	11.	Sex	0.027
3.	Mother_Height_cm	0.088	12.	Drought_Exposure	0.023
4.	Mother_BMI	0.085	13.	Marital_Status	0.021
5.	Region	0.085	14	Residence	0.020
6.	Maternal_Age	0.071	15	Food_Aid	0.020
7.	Dietary_Diversity	0.047	16	Place_of_Delivery	0.019
8	Antenatal_Visits	0.043	17	Wealth_Index	0.018
9.	Maternal_Education	0.032			

As shown in Table 4.8, child age (0.13) emerged as the strongest determinant, while wealth index (0.018) was the weakest.

4.7. Prototype

Figure 4.26 show the prototype of ensemble machine learning model for highest performance of XGBoost. This shows that the predicted status of the child based on the training data that is gathered from the ministry of health; seqota declaration and the value of threshold, if the value of threshold is greater than 0.50 indicated that the child was stunted need intervention and it displayed the factors and the solution was highlighted. Also if the probability of stunting is between 0.50 up to 0.75 it was moderate stunted and if probability of stunting is >=0.75 it was severe stunting indicates required immediate intervention. If the threshold value of the children was less than 0.50 this indicated that the child was not stunted and not required intervention.

This prototype represents an interactive prediction tool for assessing child stunting risk. It combined socioeconomic, health, and environmental variables with best ensemble machine learning model to generate individualized child predictions probability. This prototype included interface or Input Variables, Prediction Output, Threshold Control, Possible Causes and Suggested Solutions. The interface was allowed entry of key determinants or Input Variables such as child age, mother height, dietary diversity, sex, residence, drought exposure, wealth index, maternal education, sanitation, breastfeeding, conflict exposure, diarrhea, food aid, and birth weight.

Prediction Output, based on these inputs, the model was predicted whether a child is stunted or not. In the example shown, the child was predicted as Stunted with a probability of 0.83 (83%).Threshold Control**,** the slider at 0.50 determines the cutoff probability for classifying stunting. Adjusting it changes sensitivity vs. specificity. Possible Causes, the system highlights contributing factors such as low birth weight, poor dietary diversity, and Diarrhoea illness. Suggested Solutions**,** It was provided actionable recommendations such as improving maternal nutrition, promoting diverse feeding practices, and enhancing sanitation.

This prototype is not just a classifier, but it was also decision support tool**.** It bridges practical prediction with policy-relevant insights, showing both why a child is at danger and what involvements could help. It shows how ensemble machine learning could be translated into practical guidance for health workers and policymakers**,** moving beyond raw numbers to actionable strategies.

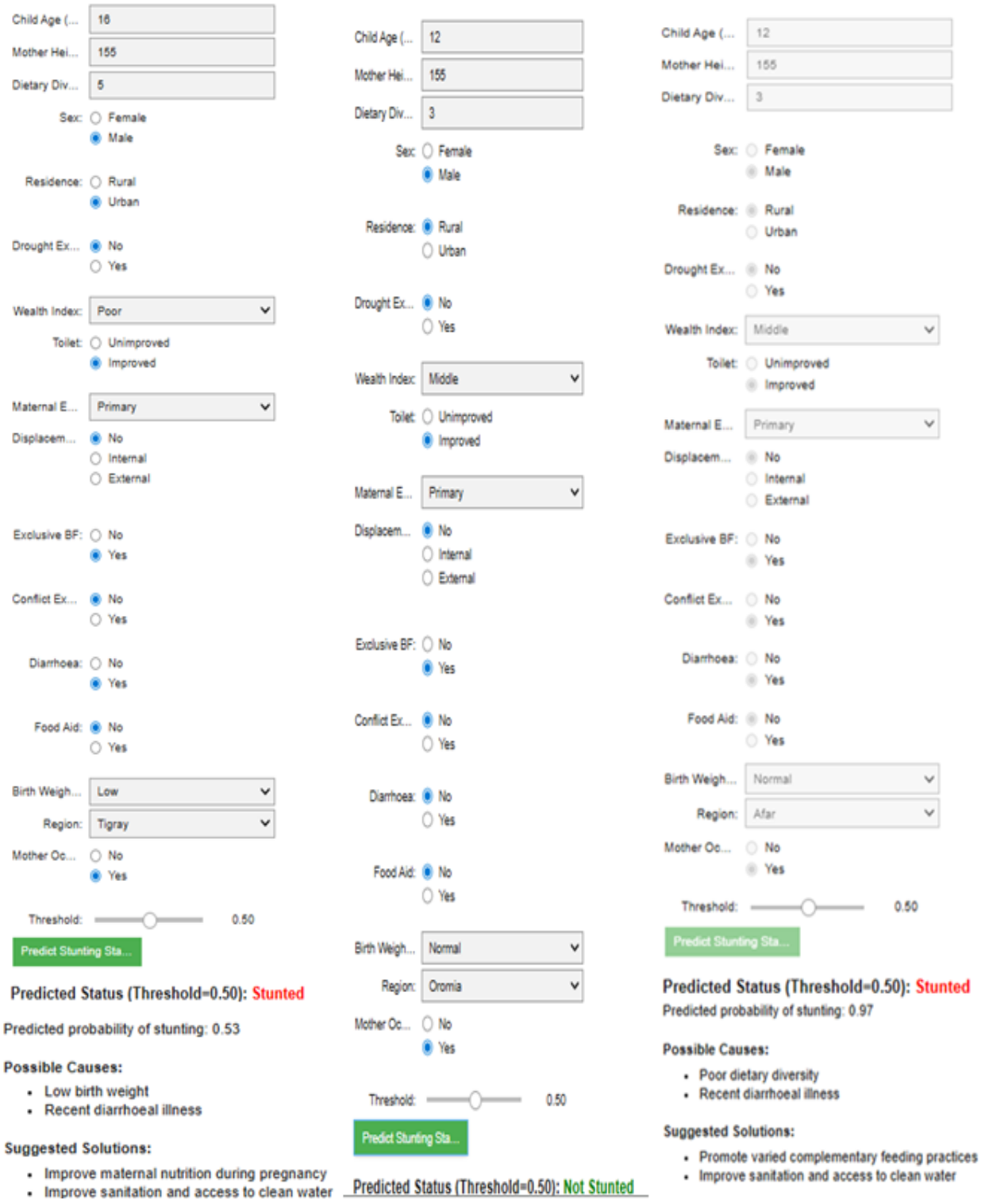


Figure4. 26: Prototype for Not Stunted and Stunted class

4.8. Feature Explainability

4.8.1. SHAP Explanation

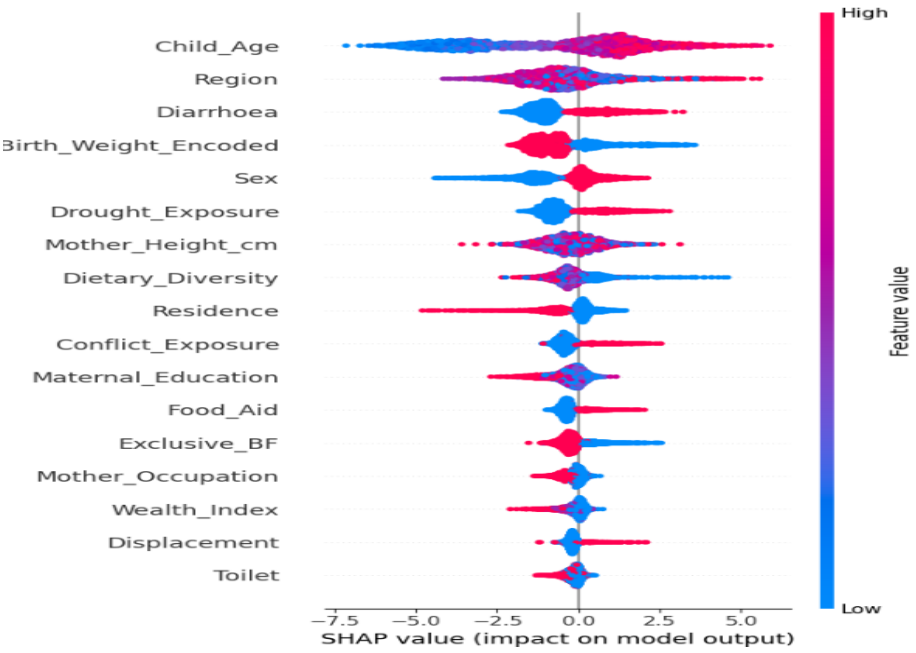


Figure4. 27: SHAP Values of Stunting Risk Factors

This plot is a SHAP (Shapley adaptive Explanation) summary plot which shows that how each feature influences the predictions of ensemble machine learning model. Child Age sits at the top, meaning it has the strongest impact on the model's output. The spread of SHAP values shows how different ages push predictions higher or lower. Region and Diarrhoea also have large effects, suggesting topographical and health conditions are key factors of stunting risk. Features such as Birth Weight, Mother Height, and Dietary Diversity indicate moderate effect, reflecting natural and nutritional factors. Social and environmental variables like Conflict Exposure, Maternal Education, and Wealth Index have smaller but still observable impacts. The blue to red color gradient specifies the feature value. Blue indicate low values like younger age, lower dietary diversity; and Red shows high values such as older age, higher dietary diversity. The x-axis or SHAP value shows whether the feature pushes the prediction toward higher risk positive SHAP or lower risk negative SHAP.

4.8.2. LIME Explanation

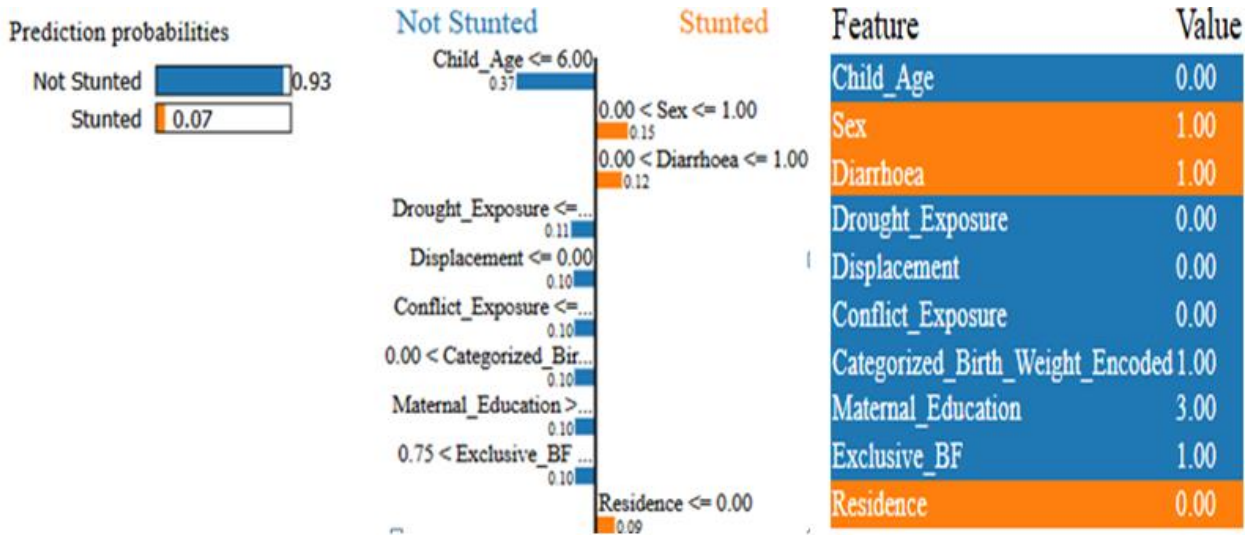


Figure4. 28: LIME Prediction Explanations

This visualization shows that prediction explanation chart for child stunting, displaying how the model reached at its decision for one child. On the left side, the bars graph show the predicted probabilities of Not Stunted is 93 % very high probability and the predicted probabilities of Stunted is 7 %(low probability), this indicates the model predicts this child is not stunted. The reasoning explains as follows, Child Age 0 months or newborn is Very young age strongly reduces stunting risk in the model, pushing the prediction toward Not Stunted. Male children coded as 1 in the dataset show slightly higher stunting risk, so this factor pushed the prediction toward Stunted. Diarrhoea present coded as 1 in the dataset show this is a risk factor, adding weight toward Stunted. Maternal Education coded as 3 in the dataset higher level (higher than the threshold) Strong protective factor, pushing the prediction toward Not Stunted. Exclusive Breastfeeding present coded as 1 in the dataset is another strong protective factor; supporting the Not Stunted outcome. Birth Weight normal coded as 1 in the dataset reduces risk, supporting Not Stunted. Environmental and Social Factors such as Drought Exposure, Conflict Exposure and Displacement All coded as absent, meaning no added risk from external shocks, but Residence rural coded as 0, sex male coded as 1 and Diarrhoea presence coded as 1 meaning have risk stunted.

In the middle section, as seen indicate the feature thresholds that influenced the model prediction. Examples Child_Age≤ 6 months pushed the prediction toward Not Stunted, Sex=1(male), Diarrhoea=1 (present) and Residence=0 (Rural) added some risk toward Stunted,

Maternal Education>3 (higher education) and Exclusive BF =1 (exclusive breastfeeding) strongly pushed toward Not Stunted, Environmental factors like Drought Exposure, Conflict Exposure, and Displacement were all coded as 0 (absence) indicates to no more risk of stunted. On the right side, the table lists shows that the actual feature values for this child such as Age = 0 months, Sex = male, Diarrhoea = yes, Maternal Education = 3, Exclusive BF = yes and so on.

4.9. Discussion

The study was designed to develop models to predict the child stunting with accuracy of 81.76% and ROC of 82% and to discuss the three research questions briefly. The 1st question of this study is "**What are the most significant risk factors used to predict for the under-two of child stunting**?" To answer the question, feature selection techniques were used from all the factors that used to develop the predictive model, and the highest contributed attributes such as Child Age, Sex, Birth weight, Mother Height, Region, Residence, Dietary Diversity, Conflict Exposure, Drought Exposure, Mother Education, Mother Occupation and Displacement are the determining of stunting based on the best performed algorithms.

The 2nd question is "**Which type of ensemble machine learning algorithm achieves the highest accuracy in predicting child stunting**?" To answer the question, five experiments for each of the five models are conducted. From those experiments XGBoosting achieved the best performance with an accuracy of 81.76%.

The 3rd question was "**What range of performed metrics did the best-performing ensemble algorithm achieve?**" To answer the research question, I had compared all the algorithms that I used in the research based on accuracy, precision, recall, F1-score and ROC values. So, the best-performing ensemble algorithm XGBoosting achieved performance in the range of 80.79%–82.81% across F1-score, precision, accuracy, recall and ROC-AUC.

Based on these results, I can conclude that the proposed predictive models, particularly XGBoost and Cat Boost demonstrate relatively good performance in accurately identifying stunting status and the rests are relatively poor performance.

4.9.1. Direct comparison with Ethiopian studies

Comparison with Ayele et al. 2025 and Wudu 2026

Table4. 9: comparison to related study

Characteristics	Ayele et al. 2025	Wudu 2026	Current study	Interpretation
Age group	under five	under five	Under two	Current study Targets reversible window
Date	EDHS 2011-2016	EDHS 2019	seqota 2021-2026	current study use recent, program specific data
Accuracy	98%	94%	81.8%	Low but on a harder task (prediction vs classification)
AUC	99.9%	---	82.8%	Ayele's AUC likely overstated by data leakage using post 24 month
Feature	28 EDHS variables	21 EDHS variables	27 including contextual shocks	current study was Inclusive
Interpretability	None	LIME	SHAP and LIME	current study enables policy change and intervention
Prototype	None	None	Yes	predict and identify further risk children

1.Age specific difficulty: predicting stunting in under two children before irreversible damage is inherently more challenging than classifying stunting in under five children where stunting Is already established. The current study attempts on prediction where as Ayele (2025) and Wudu (2026) performs classification.

2. Recent Data: current study used recent data and includes contextual factors while, Ayele and Wudu's data precedes those major shocks.

3. Interpretability loose: SHAP and LIME indicates why prediction is made and enabling targeted intervention. Black-box models achieving 98% and 94 % accuracy respectively but, cannot guide policy. LIME explainability with Hybrid model is not appropriated and accuracy alone is insufficient in Wudu (2026) study.

4. Clinical utility: a prototype with threshold sliders and actionable recommendations has greater real-world impact than accuracy alone.

There are many key innovation enabling improvement such as Boruta feature selection vs correlation based, optuna hyperparameter optimization vs default parameters, inclusion of contextual factors to capture additional signal and under two focuses to reduce heterogeneity from older children.

CHAPTER FIVE: CONCLUSION AND RECOMMENDATIONS

5.1. Conclusion

This study developed a predicting model for stunting among under-two children using Ensemble machine learning approaches for Ethiopian children dataset sourced from the ministry of health conducted between 2021 and 2026. It compares the predictive performance of various ensemble machines learning model integrating with the main risk factors, SMOTE and hyperparameters optimization Techniques. The results show that the Extreme Gradient Boosting and Cat Boost models performed significantly better than the other models with high accuracy and ROC-AUC scores. The study also identified the basic predictors of stunting such as Child Age, Birth weight, Mother Height, Mother BMI and overall socioeconomic status. The findings show that early identification of stunted children was feasible, and enabling timely interventions during the critical period of time, enhance the understanding of stunting in Ethiopian and provide a basis for more targeted study and actions aimed at lowering the rates of stunting.

The current study sacrifices some predictive accuracy (compared to potentially over fired prior work) in exchange of for interpretability, contextual relevance, and clinical utility representing an appropriate trade-off for policy relevant research.

5.2. Strengths and Limitations

Concerning the strength of this study, intended prediction accuracy was achieved. Prediction of stunting with good predictive performance ensures the potential capability of ensemble machine learning techniques in the field of health care such as child stunting. Additional, since the objectives of this research were achieved, it can be concluded that the study was successful overall with an accuracy of 81.6%. However, the main limitation of the research is that the quality of the dataset is often complex due to the heterogeneity and no external validation from other dataset.

5.3. Contributions

The contributions of this research work are Constructed a classification of stunting and prediction model by including, identify additional risk factors and socioeconomic level for stunting, used recent data, focused on reversible children age group and make model

explainability using local and global explainability tools for both policy maker and health experts to make intervention by any governmental and private sectors.

5.4. Recommendation

This study forwards the following recommendations for further investigation based on the research's findings. The study is dependent on ministry of health from 2021 up to 2026. To enhance the predictive performance of the current study should explore additional attributes, more recent data from other institution, alternative class balanced techniques, advanced algorithms such as deep learning via sophisticated feature selection technique and hyper parameter tuning and mobile app based prototype for field use could be areas of exploration for future work.

REFERENCE

[1] M. S. Islam, A. N. Zafar Ullah, S. Mainali, M. A. Imam, and M. I. Hasan, "Determinants of stunting during the first 1,000 days of life in Bangladesh: A review," *Food Sci. Nutr.*, vol. 8, no. 9, pp. 4685–4695, 2020, doi: 10.1002/fsn3.1795.

[2] S. Ndagijimana, I. H. Kabano, E. Masabo, and J. M. Ntaganda, "Prediction of Stunting among Under-5 Children in Rwanda Using Machine Learning Techniques," *J. Prev. Med. Public Heal.*, vol. 56, no. 1, pp. 41–49, 2023, doi: 10.3961/jpmph.22.388.

[3] UNICEF/WHO/WORLD BANK, "Levels and trends in child malnutrition UNICEF / WHO / World Bank Group Joint Child Malnutrition Estimates Key findings of the 2021 edition," *World Heal. Organ.*, pp. 1–32, 2021, [Online]. Available: https://www.who.int/publications/i/item/9789240025257

[4] B. Zegeye, G. Shibre, D. Idriss-Wheeler, and S. Yaya, "Trends in inequalities in childhood stunting in Ethiopia from 2000 to 2016: A cross sectional study," *J. Public Heal. (United Kingdom)*, vol. 43, no. 3, pp. 655–663, 2021, doi: 10.1093/pubmed/fdaa051.

[5] Y. T. Haile and T. Z. Legisso, "Trends and determinants of child stunting in Ethiopia: a longitudinal analysis of sociodemographic factors and regional disparities," *BMC Public Health*, vol. 24, no. 1, 2024, doi: 10.1186/s12889-024-20962-3.

[6] R. Maulina, M. B. Qomaruddin, B. Prasetyo, R. Indawati, and R. Alfitri, "The Effect of Stunting on the Cognitive Development in Children: A Systematic Review and Meta-analysis," *Stud. Ethno-Medicine*, vol. 17, no. 1–2, pp. 19–27, 2023, doi: 10.31901/24566772.2023/17.1-2.661.

[7] C. Berti and P. Socha, "Infant and Young Child Feeding Practices and Health," *Nutrients*, vol. 15, no. 5, pp. 13–17, 2023, doi: 10.3390/nu15051184.

[8] M. G. Jebena *et al.*, "Barriers and facilitators of maternal health care services use among pastoralist women in Ethiopia: Systems thinking perspective," *Pastoralism*, vol. 12, no. 1, 2022, doi: 10.1186/s13570-022-00236-6.

[9] Ministry of Health of Ethiopia, "Seqota Declaration," *Seqota Declar.*, pp. 1–15, 2015.

[10] T. Gebreegziabher and N. Regassa, "Ethiopia's high childhood undernutrition explained: Analysis of the prevalence and key correlates based on recent nationally representative data," *Public Health Nutr.*, vol. 22, no. 11, pp. 2099–2109, 2019, doi: 10.1017/S1368980019000569.

[11] H. Ashar *et al.*, "Factors related to stunting in children under 2 years old in the Papua, Indonesia Does the type of residence matter?," *Saudi Med. J.*, vol. 45, no. 3, pp. 273–278, 2024, doi: 10.15537/smj.2024.45.3.20230774.

[12] T. T. Tadele, C. C. Gebremedhin, M. U. Markos, and E. L. Fitsum, "Stunting and associated factors among 6–23 month old children in drought vulnerable kebeles of Demba Gofa district, southern Ethiopia," *BMC Nutr.*, vol. 8, no. 1, pp. 1–11, 2022, doi:

10.1186/s40795-022-00501-2.

[13] F. H. Bitew, C. S. Sparks, and S. H. Nyarko, “Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia,” *Public Health Nutr.*, vol. 25, no. 2, pp. 269–280, 2022, doi: 10.1017/S1368980021004262.

[14] A. Byna, F. N. Anisa, and N. Nurhaeni, “Improving Stunting Prediction in Children: Evaluating Ensemble Algorithms with SMOTE and Feature Selection,” *AIP Conf. Proc.*, vol. 3250, no. 1, 2025, doi: 10.1063/5.0240617.

[15] M. K. Ayele, G. A. Baye, S. H. Yesuf, A. A. Engda, and E. T. Mitiku, “Predicting stunting status among under five children in ethiopia using ensemblemachine learning algorithms,” 2025. doi: 10.1038/s41598-025-03206-1.

[16] UNICEF, WHO, and W. B. Group, “Levels and trends in child malnutrition: Key finding of the 2023 edition,” *Asia-Pacific Popul. J.*, vol. 24, no. 2, pp. 51–78, 2023.

[17] B. L. Seifu, G. A. Tesema, B. M. Fentie, T. Z. Yehuala, A. H. Moloro, and K. U. Mare, “Geographical variation in hotspots of stunting among under-five children in Ethiopia: A geographically weighted regression and multilevel robust Poisson regression analysis,” *PLoS One*, vol. 19, no. 5 May, pp. 1–23, 2024, doi: 10.1371/journal.pone.0303071.

[18] FMoH, “Government of Ethiopia Commitment Goals and Actions 2026 – 2030,” no. March 2025, 2026.

[19] U. WHO, “Global nutrition targets 2025: breastfeeding policy brief (WHO/NMH/NHD14. 7),” *Geneva World Heal. Organ.*, 2014.

[20] Z. Mei and L. M. Grummer-Strawn, “Standard deviation of anthropometric Z-scores as a data quality assessment tool using the 2006 WHO growth standards: A cross country analysis,” *Bull. World Health Organ.*, vol. 85, no. 6, pp. 441–448, 2007, doi: 10.2471/BLT.06.034421.

[21] UNICEF, “Conceptual framework of young child nutrition,” 1990.

[22] J. Hoddinott, H. Alderman, J. R. Behrman, L. Haddad, and S. Horton, “The economic rationale for investing in stunting reduction,” *Matern. Child Nutr.*, vol. 9, no. S2, pp. 69–82, 2013, doi: 10.1111/mcn.12080.

[23] E. Mulu and B. Mengistie, “Household food insecurity and its association with nutritional status of under five children in Sekela District, Western Ethiopia: A comparative cross-sectional study,” *BMC Nutr.*, vol. 3, no. 1, pp. 1–9, 2017, doi: 10.1186/s40795-017-0149-z.

[24] A. J. Prendergast and J. H. Humphrey, “The stunting syndrome in developing countries,” *Paediatr. Int. Child Health*, vol. 34, no. 4, pp. 250–265, 2014, doi: 10.1179/2046905514Y.0000000158.

[25] G. A. Lemlem, M. K. Mezen, A. Atinafu, and Z. A. Abitew, “Maternal factors associated with low birth weight in governmental hospitals of Wollo District, Northeast Ethiopia: a

cross sectional study," *Pan African Med. J. One Heal.*, vol. 4, no. 18, 2021, doi: 10.11604/pamjoh.2021.4.18.27861.

[26] A. C. Miller, M. B. Murray, D. R. Thomson, and M. C. Arbour, "How consistent are associations between stunting and child development? Evidence from a meta-analysis of associations between stunting and multidimensional child development in fifteen low- and middle-income countries," *Public Health Nutr.*, vol. 19, no. 8, pp. 1339–1347, 2016, doi: 10.1017/S136898001500227X.

[27] A. Wemakor, H. Garti, T. Azongo, H. Garti, and A. Atosona, "Young maternal age is a risk factor for child undernutrition in Tamale Metropolis, Ghana," *BMC Res. Notes*, vol. 11, no. 1, pp. 1–5, 2018, doi: 10.1186/s13104-018-3980-7.

[28] A. D. Laksono, R. D. Wulandari, N. Amaliah, and R. W. Wisnuwardani, "Stunting among children under two years in Indonesia: Does maternal education matter?," *PLoS One*, vol. 17, no. 7 July, pp. 1–9, 2022, doi: 10.1371/journal.pone.0271509.

[29] M. W. Cooper *et al.*, "Mapping the effects of drought on child stunting," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 35, pp. 17219–17224, 2019, doi: 10.1073/pnas.1905228116.

[30] D. K. Kinyoki *et al.*, "Conflict in Somalia: Impact on child undernutrition," *BMJ Glob. Heal.*, vol. 2, no. 2, 2017, doi: 10.1136/bmjgh-2016-000262.

[31] M. M. Azanaw, D. T. Anley, R. M. Anteneh, G. Arage, and A. A. Muche, "Effects of armed conflicts on childhood undernutrition in Africa: a systematic review and meta-analysis," *Syst. Rev.*, vol. 12, no. 1, pp. 1–10, 2023, doi: 10.1186/s13643-023-02206-4.

[32] L. Cao, "Data science: A comprehensive overview," *ACM Comput. Surv.*, vol. 50, no. 3, 2018, doi: 10.1145/3076253.

[33] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, p. 160, 2021, doi: 10.1007/s42979-021-00592-x.

[34] M. Chen, S. Mangalathu, and J.-S. Jeon, "Machine Learning–Based Seismic Reliability Assessment of Bridge Networks," *J. Struct. Eng.*, vol. 148, no. 7, pp. 1–4, 2022, doi: 10.1061/(asce)st.1943-541x.0003376.

[35] S. Razia, P. S. Prathyusha, N. V. Krishna, and N. S. Sumana, "A Comparative study of machine learning algorithms on thyroid disease prediction," *Int. J. Eng. Technol*, vol. 7, no. 2.8, pp. 315–319, 2018.

[36] I. Iswanto, K. Shankar, P. T. Nguyen, W. Hashim, and A. Maseleno, "Identifying diseases and diagnosis using machine learning," *All Open Access, Bronze*, 2019.

[37] I. M. Ibrahim and A. M. Abdulazeez, "The Role of Machine Learning Algorithms for Diagnosing Diseases," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 1, pp. 10–19, 2021, doi: 10.38094/jastt20179.

[38] S. Asif *et al., Advancements and Prospects of Machine Learning in Medical Diagnostics:*

Unveiling the Future of Diagnostic Precision, vol. 32, no. 2. 2025. doi: 10.1007/s11831-024-10148-w.

[39] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput. Sci.*, vol. 2, no. 3, pp. 1–26, 2021, doi: 10.1007/s42979-021-00592-x.

[40] C. Song, T. Ristenpart, and V. Shmatikov, "Machine learning models that remember too much," *Proc. ACM Conf. Comput. Commun. Secur.*, no. December, pp. 587–601, 2017, doi: 10.1145/3133956.3134077.

[41] M. S. Haris, M. Anshori, and A. N. Khudori, "Prediction of Stunting Prevalence in East Java Province With Random Forest Algorithm," *J. Tek. Inform.*, vol. 4, no. 1, pp. 11–13, 2023, doi: 10.52436/1.jutif.2023.4.1.614.

[42] N. S. O'Connell, B. C. Jaeger, G. S. Bullock, and J. L. Speiser, "A comparison of random forest variable selection methods for regression modeling of continuous outcomes," *Brief. Bioinform.*, vol. 26, no. 2, pp. 1–7, 2025, doi: 10.1093/bib/bbaf096.

[43] N. Khasanah, R. Komarudin, N. Afni, Y. I. Maulana, and A. Salim, "Skin Cancer Classification Using Random Forest Algorithm," *Sisfotenika*, vol. 11, no. 2, p. 137, 2021, doi: 10.30700/jst.v11i2.1122.

[44] L. Wan, K. Gong, G. Zhang, X. Yuan, C. Li, and X. Deng, "An efficient rolling bearing fault diagnosis method based on spark and improved random forest algorithm," *IEEE Access*, vol. 9, pp. 37866–37882, 2021, doi: 10.1109/ACCESS.2021.3063929.

[45] C. Song, T. Ristenpart, and V. Shmatikov, "Machine learning models that remember too much," in *Proceedings of the 2017 ACM SIGSAC Conference on computer and communications security*, 2017, pp. 587–601.

[46] D. James, *Introduction to Machine Learning with Python: A Guide for Beginners in Data Science*. CreateSpace Independent Publishing Platform, 2018.

[47] B. Naderalvojoud and T. Hernandez-Boussard, "Improving machine learning with ensemble learning on observational healthcare data," *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2023, pp. 521–529, 2023.

[48] I. D. Mienye and Y. Sun, "A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects," *IEEE Access*, vol. 10, no. August, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.

[49] H. Rao *et al.*, "Feature selection based on artificial bee colony and gradient boosting decision tree," *Appl. Soft Comput. J.*, vol. 74, no. January, pp. 634–642, 2019, doi: 10.1016/j.asoc.2018.10.036.

[50] W. Liang, S. Luo, G. Zhao, and H. Wu, "Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms," *Mathematics*, vol. 8, no. 5, pp. 1–17, 2020, doi: 10.3390/MATH8050765.

[51] J. Nobre and R. F. Neves, “Combining principal component analysis, discrete wavelet transform and XGBoost to trade in the financial markets,” *Expert Syst. Appl.*, vol. 125, pp. 181–194, 2019.

[52] B. Zhang, Y. Zhang, and X. Jiang, “Feature selection for global tropospheric ozone prediction based on the BO-XGBoost-RFE algorithm,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–10, 2022, doi: 10.1038/s41598-022-13498-2.

[53] D. A. McCarty, H. W. Kim, and H. K. Lee, “Evaluation of light gradient boosted machine learning technique in large scale land use and land cover classification,” *Environ. - MDPI*, vol. 7, no. 10, pp. 1–22, 2020, doi: 10.3390/environments7100084.

[54] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” *Adv. Neural Inf. Process. Syst.*, vol. 31, 2018.

[55] J. Tanha, Y. Abdi, N. Samadi, N. Razzaghi, and M. Asadpour, “Boosting methods for multi-class imbalanced data classification: an experimental review,” *J. Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00349-y.

[56] A. V. Dorogush, V. Ershov, and A. Gulin, “CatBoost: gradient boosting with categorical features support,” pp. 1–7, 2018, [Online]. Available: http://arxiv.org/abs/1810.11363

[57] T. Yu and H. Zhu, “Hyper-parameter optimization: A review of algorithms and applications,” *arXiv Prepr. arXiv2003.05689*, 2020.

[58] A. Kumar, M. Nadeem, and H. Banka, “Nature inspired optimization algorithms: a comprehensive overview,” *Evol. Syst.*, vol. 14, no. 1, pp. 141–156, 2023.

[59] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization,” *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, 2012.

[60] T. Domhan, J. T. Springenberg, and F. Hutter, “Speeding up automatic hyperparameter optimization of deep neural networks by extrapolation of learning curves,” *IJCAI Int. Jt. Conf. Artif. Intell.*, vol. 2015-Janua, pp. 3460–3468, 2015.

[61] S. Shekhar, A. Bansode, and A. Salim, “A Comparative study of Hyper-Parameter Optimization Tools,” *2021 IEEE Asia-Pacific Conf. Comput. Sci. Data Eng. CSDE 2021*, pp. 11–14, 2021, doi: 10.1109/CSDE53843.2021.9718485.

[62] K. Yang, L. Liu, and Y. Wen, “The impact of Bayesian optimization on feature selection,” *Sci. Rep.*, vol. 14, no. 1, pp. 1–34, 2024, doi: 10.1038/s41598-024-54515-w.

[63] F. O. R. F. Code and C. Here, “Bayesian Hyperparameter Tuning with Optuna : A Beginner-Friendly Guide,” pp. 1–22, 2025.

[64] B. A. Dada, N. I. Nwulu, and S. O. Olukanmi, “Bayesian optimization with Optuna for enhanced soil nutrient prediction: a comparative study with genetic algorithm and particle swarm optimization,” *Smart Agric. Technol.*, vol. 12, no. October 2024, p. 101136, 2025, doi: 10.1016/j.atech.2025.101136.

[65] U. Bhatt *et al.*, “Explainable machine learning in deployment,” in *Proceedings of the 2020 conference on fairness, accountability, and transparency*, 2020, pp. 648–657.

[66] R. O. Alabi, M. Elmusrati, I. Leivo, A. Almangush, and A. A. Mäkitie, “Machine learning explainability in nasopharyngeal cancer survival using LIME and SHAP,” *Sci. Rep.*, vol. 13, no. 1, pp. 1–14, 2023, doi: 10.1038/s41598-023-35795-0.

[67] R. Alkhanbouli, H. Matar Abdulla Almadhaani, F. Alhosani, and M. C. E. Simsekler, “The role of explainable artificial intelligence in disease prediction: a systematic literature review and future research directions,” *BMC Med. Inform. Decis. Mak.*, vol. 25, no. 1, 2025, doi: 10.1186/s12911-025-02944-6.

[68] O. N. Chilyabanyama *et al.*, “Performance of Machine Learning Classifiers in Classifying Stunting among Under-Five Children in Zambia,” *Children*, vol. 9, no. 7, 2022, doi: 10.3390/children9071082.

[69] M. A. E. Pratama, S. Hendra, H. R. Ngemba, R. Nur, R. Azhar, and R. Laila, “Comparison of Machine Learning Algorithms for Predicting Stunting Prevalence in Indonesia,” *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 13, no. 2, pp. 200–209, 2024, doi: 10.32736/sisfokom.v13i2.2097.

[70] H. Shen, H. Zhao, and Y. Jiang, “Machine Learning Algorithms for Predicting Stunting among Under-Five Children in Papua New Guinea,” *Children*, vol. 10, no. 10, 2023, doi: 10.3390/children10101638.

[71] B. Fogarty, A. García-Martínez, N. V. Chawla, and E. Serván-Mori, “Social and economic predictors of under-five stunting in Mexico: a comprehensive approach through the XGB model,” *J. Glob. Health*, vol. 15, p. 04065, 2025, doi: 10.7189/jogh.15.04065.

[72] S. M. J. Rahman *et al.*, “Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach,” *PLoS One*, vol. 16, no. 6 June 2021, pp. 1–11, 2021, doi: 10.1371/journal.pone.0253172.

[73] A. B. Zemariam *et al.*, “Prediction of stunting and its socioeconomic determinants among adolescent girls in Ethiopia using machine learning algorithms,” 2025. doi: 10.1371/journal.pone.0316452.

[74] A. Hendy *et al.*, “Supervised machine learning for classification and prediction of stunting among under-five Egyptian children,” *BMC Pediatr.*, vol. 25, no. 1, 2025, doi: 10.1186/s12887-025-06138-x.

[75] T. K. Wudu, A. A. Endalew, and A. A. Dires, “Explainable hybrid machine learning model for predicting stunting and identifying key risk factors among Ethiopian children under five,” pp. 1–17, 2026.

[76] A. M. Messele, “Ensemble machine learning for predicting academic performance in STEM education,” *Discov. Educ.*, vol. 4, no. 1, 2025, doi: 10.1007/s44217-025-00710-4.

[77] A. M. Sharifnia, D. E. Kpormegbey, D. K. Thapa, and M. Cleary, “A Primer of Data

Cleaning in Quantitative Research: Handling Missing Values and Outliers," *J. Adv. Nurs.*, vol. 82, no. 1, pp. 970–975, 2026, doi: 10.1111/jan.16908.

[78] A. P. Joshi and B. V. Patel, "Data Preprocessing: The Techniques for Preparing Clean and Quality Data for Data Analytics Process," *Orient. J. Comput. Sci. Technol.*, vol. 13, no. 0203, pp. 78–81, 2021, doi: 10.13005/ojcst13.0203.03.

[79] S. Borrohou, R. Fissoune, and H. Badir, "The role of data transformation in modern analytics: A comprehensive survey," *J. Comput. Lang.*, vol. 84, no. September, pp. 4–9, 2025, doi: 10.1016/j.cola.2025.101329.

[80] E. Toulabinejad, M. Mirsafaei, and A. Basiri, "Supervised discretization of continuous-valued attributes for classification using RACER algorithm," *Expert Syst. Appl.*, vol. 244, no. June, pp. 1–7, 2024, doi: 10.1016/j.eswa.2023.121203.

[81] V. V. Starovoitov and Y. I. Golub, "Data normalization in machine learning," *Informatics*, vol. 18, no. 3, pp. 83–96, 2021, doi: 10.37661/1816-0301-2021-18-3-83-96.

[82] P. Agrawal, H. F. Abutarboush, T. Ganesh, and A. W. Mohamed, "Metaheuristic algorithms on feature selection: A survey of one decade of research (2009-2019)," *IEEE Access*, vol. 9, pp. 26766–26791, 2021, doi: 10.1109/ACCESS.2021.3056407.

[83] G. Manikandan, B. Pragadeesh, V. Manojkumar, A. L. Karthikeyan, R. Manikandan, and A. H. Gandomi, "Classification models combined with Boruta feature selection for heart disease prediction," *Informatics Med. Unlocked*, vol. 44, pp. 1–33, 2024, doi: 10.1016/j.imu.2023.101442.

[84] P. Soltanzadeh and M. Hashemzadeh, "RCSMOTE: Range-Controlled synthetic minority over-sampling technique for handling the class imbalance problem," *Inf. Sci. (Ny).*, vol. 542, no. January, pp. 92–111, 2021, doi: 10.1016/j.ins.2020.07.014.

[85] Muraina and I. Olaniyi, "Ideal Dataset Splitting Ratios in Machine Learning Algorithms :," *7th Int. Mardin Artuklu Sci. Res. Conf.*, no. February, pp. 496–504, 2022.

APPENDICES

Appendix A:

Source Code Snippet

```
# HAZ removed to prevent data leakage (outcome variable not used as
predictor)
if 'HAZ' in df.columns:
    df = df.drop(columns=["HAZ"])
# Handling missing values using mean and mode imputation
print(f"Missing values in 'Region' before imputation:
{df['Region'].isnull().sum()}")

# Calculate the mode of the 'Mother_Height_cm' column
mode_Region = df['Region'].mode()[0]

# Fill missing values with the mode
df['Region'] = df['Region'].fillna(mode_Region)

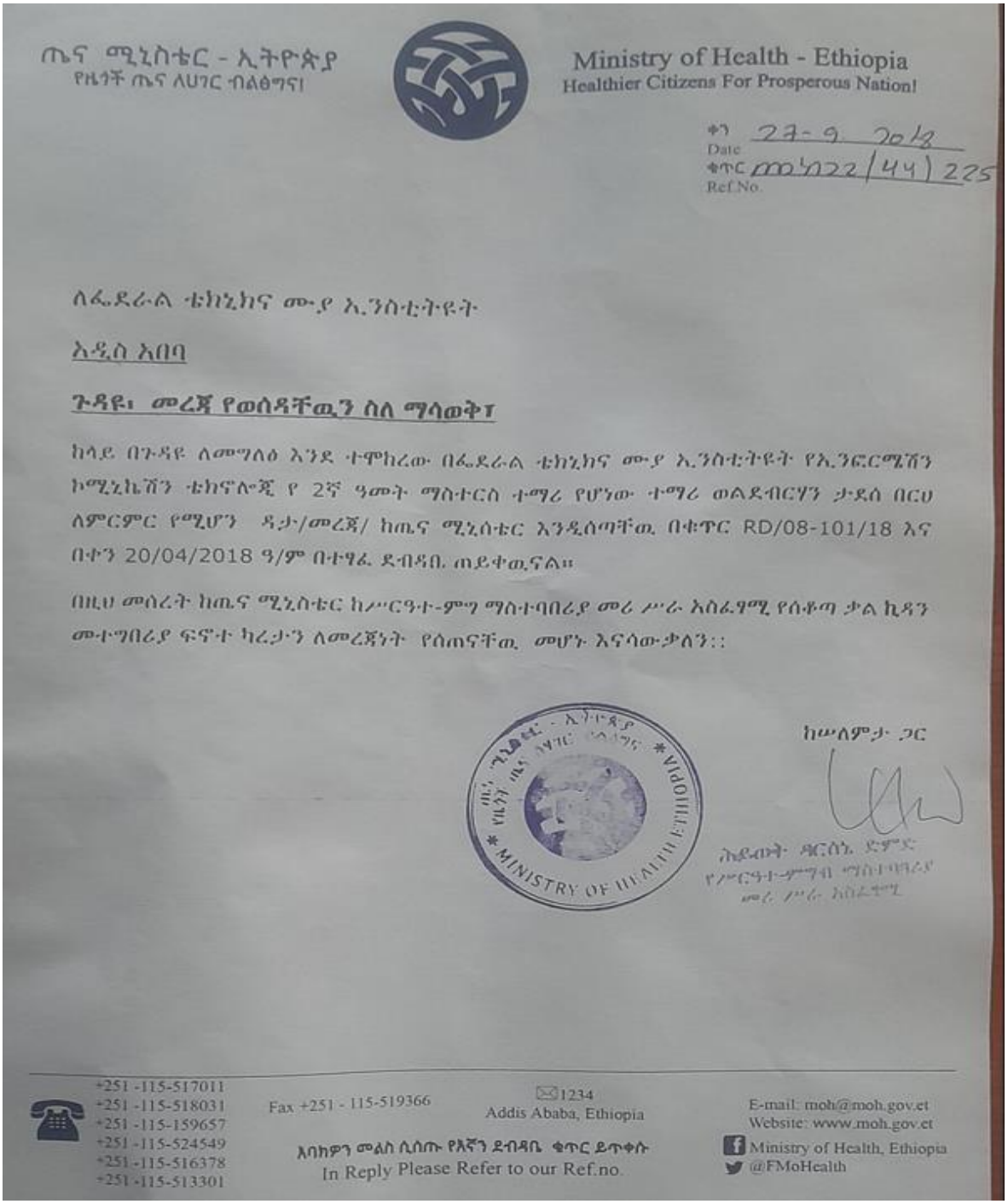
print(f"Missing values in 'Region' after imputation:
{df['Region'].isnull().sum()}")

#Feature explainability  using SHAP and LIME
import shap
import numpy as np
import pandas as pd

explainer = shap.TreeExplainer(final_lgbm_model['model'])
# Calculate SHAP values for the test set
shap_values = explainer.shap_values(X_test_eval)
# Plot the SHAP summary plot
print("Generating SHAP summary plot...")
shap.summary_plot(shap_values, X_test_eval, show=False)
import matplotlib.pyplot as plt
plt.show()
from lime.lime_tabular import LimeTabularExplainer
import numpy as np
import pandas as pd
class_names = ['Not Stunted', 'Stunted'] # Adjust if your target
labels are different
# Initialize the LIME explainer
lime_explainer = LimeTabularExplainer(
    training_data=np.array(X_train_eval), # Use training data
for the explainer
```

```
    feature_names=X_selected.columns.tolist(), # Use actual
feature names
    class_names=class_names,
    mode="classification"
)
# Choose an instance from the test set to explain (e.g., the 5th
instance)
i = 5
instance_to_explain = X_test_eval.iloc[i].values
# Generate explanation for the chosen instance
print(f"Generating LIME explanation for instance {i} from the
test set...")
exp = lime_explainer.explain_instance(
    instance_to_explain,
    final_lgbm_model.predict_proba # Use the predict_proba
method of the pipeline
)
# Display the LIME explanation in the notebook
exp.show_in_notebook(show_all=False)
```

Appendices B: approval Letter for data collection from Ministry of Health



ጤና ሚኒስቴር - ኢትዮጵያ
የዜጎች ጤና ለሀገር ብልፅግና!

Ministry of Health - Ethiopia
Healthier Citizens For Prosperous Nation!

ቀን 27-9-2010
Date
ቁጥር መክ22/44)225
Ref.No.

ለፌደራል ቴክኒክና ሙያ ኢንስቲትዩት

አዲስ አበባ

ጉዳዩ፣ መረጃ የወሰዳቸዉን ስለ ማሳወቅ፣

ከላይ በጉዳዩ ለመግለፅ እንደ ተሞከረው በፌደራል ቴክኒክና ሙያ ኢንስቲትዩት የኢንፎርሜሽን ኮሚኒኬሽን ቴክኖሎጂ የ 2ኛ ዓመት ማስተርስ ተማሪ የሆነው ተማሪ ወልደብርሃን ታደሰ በርሀ ለምርምር የሚሆን ዳታ/መረጃ/ ከጤና ሚኒሰቴር እንዲሰጣቸዉ በቁጥር RD/08-101/18 እና በቀን 20/04/2018 ዓ/ም በተፃፈ ደብዳቤ ጠይቀዉናል።

በዚህ መሰረት ከጤና ሚኒስቴር ከሥርዓተ-ምግ ማስተባበሪያ መሪ ሥራ አስፈፃሚ የሰጡት ቃል ኪዳን መተግበሪያ ፍኖተ ካረታን ለመረጃነት የሰጠናቸዉ መሆኑ እናሳውቃለን::

ከሠላምታ ጋር

ሔደሙት ጳውሰኔ ድምር
የሥርዓተ-ምግብ ማስተባበሪያ
መሪ ሥራ አስፈፃሚ

+251 -115-517011
+251 -115-518031
+251 -115-159657
+251 -115-524549
+251 -115-516378
+251 -115-513301

Fax +251 - 115-519366

1234
Addis Ababa, Ethiopia

E-mail: moh@moh.gov.et
Website: www.moh.gov.et
Ministry of Health, Ethiopia
@FMoHealth

እባክዎን መልስ ሲሰጡ የእኛን ደብዳቤ ቁጥር ይጥቀሱ
In Reply Please Refer to our Ref.no.