Machine Learning Classifier Algorithms for Predicting Malnutrition Among Under Five Children in Asia.

By:

Md. Arafat Islam

ID: 201-15-3672

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering.

Supervised By:

Ms. Amatul Bushra Akhi

Assistant professor

Department of CSE

Daffodil International University

Co-Supervised By:

Mohmmaad Asifur Rahim

Lecturer

Department of CSE

Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
DHAKA, BANGLADESH

JANUARY 2025

APPROVAL

This thesis titled "**Machine learning classifier algorithms for predicting malnutrition among under five children in Asia"** submitted by **Md. Arafat Islam (ID: 201-15-3672)** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

Dr. Sheak Rashed Haider Noori (SRH) — **Chairman**
Professor and Head Chairman
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Sharmin Akter (SNA) — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Mr. Md Mohammad Masum Bakaul (MB) — **Internal Examiner**
Sr. Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Md. Zulfiker Mahmud — **External Examiner**
Professor
Department of Computer Science and Engineering
Jagannath University

i ©Daffodil International University

DECLARATION

It hereby declares that this thesis has been done by me under the supervision of **Ms. Amatul Bushra Akhi (Assistant Professor),** Department of Computer Science and Engineering, Daffodil International University. It also declares that neither this thesis nor any part of this has been submitted anywhere else for award of any degree.

Supervised by:

Ms. Amatul Bushra Akhi

Assistant Professor

Department of Computer Science and Engineering

Daffodil International University

Co-Supervised by:

Mohmmaad Asifur Rahim

Lecturer

Department of Computer Science and Engineering

Daffodil International University

Submitted By:

Name: Md. Arafat Islam

ID: 201-15-3672

Batch: 55

Department of Computer Science and Engineering

Daffodil International University

ii ©Daffodil International University

ACKNOWLEDGEMENT

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has as-sisted us in one way or another.

First I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes it possible for us to complete the final thesis successfully.

I am really grateful and wish our profound indebtedness to Ms. Amatul Bushra Akhi, Assistant Professor, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “MACHINE LEARNING” to carry out this thesis. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior drafts and correcting them at all stages have made it possible to complete this thesis.

I would like to express our heartiest gratitude to Professor Dr. Syed Akhter Hossain, Professor, and Head, Department of CSE, for his kind help to finish our thesis and also to other faculty members and the staff of CSE department of Daffodil International University.

I would like to thank my entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and passion of my parents.

ABSTRACT

Malnourished children may have serious health issues. Furthermore, doctors often struggle to pinpoint the underlying causes of their patients' ailments, leading them to perform surgeries that may not be appropriate for all children. This is a frequent reason why children die. As a result, undernourished children are put in grave danger. Therefore, the primary goal of our research is to use AI to forecast the starvation status of children aged 0 to 5 in Asia. We looked for active research papers from 2010 to 2020 that accepted our point of view, consolidated the data, and attempted to identify benefits and downsides. Like I said before, we used an acceptable open-source dataset for this. They also studied several articles to gain an understanding of the benefits and drawbacks of ML techniques. Eight common ML classifiers Random Forest, Support Vector Machine, K-Nearest Neighbors, Logistic Regression, Bernolli Naive Bayes, Complement Naive Bayes, Decision Tree, and Gradient Boosting predict malnutrition in children under 5 with excellent accuracy. Finally, they searched for algorithms with the highest accuracy scores. Logistic Regression and K-Nearest Neighbors performed best, with train accuracy of 1.000 and 0.98 and success rates of 95.34% and 93.02%, respectively. Furthermore, the application of logistic regression classification indicated a very significant capacity to detect differences. They looked at eight machine learning algorithms to discover which one was the most successful. Among them, Logistic Regression and K-Nearest Neighbors do very well. My aim is to alleviate the future suffering of malnourished children. My next research will focus on Bangladesh's highland and coastal areas, which have poor educational levels and a high risk of child marriage.

TABLE OF CONTENTS

CONTENTS	PAGE

LIST OF FIGURES

FIGURE	PAGE

LIST OF TABLE

TABLE	PAGE

CHAPTER 1

INTRODUCTION

1.1 Introduction:

When someone's diet is deficient in certain nutrients or their body is unable to absorb the nutrients from their food, it may lead to malnutrition. Food insecurity includes deficiencies in essential nutrients as well as excess nutrient consumption, which can lead to obesity and food-related diseases. The causes of malnutrition in children are many. Insufficient protein and vitamin intake is a main reason for starvation. In most circumstances, the infant will not be able to obtain adequate protein since the diet lacks sufficient amounts of essential foods such as meat, eggs, and fish [1]. Malnutrition may result from eating too little, eating a toxic diet, or having health issues that prevent you from using nutrients properly. This disease is characterized by a poor dietary intake of calories and other nutrients. Poor diet or nutritional absorption may cause malnutrition. Undernourishment may be a side effect of cancer therapy. For decades, researchers have had a lot of interest in studying under-five children's malnutrition. Nearly all researchers have already examined the many causes of malnutrition in children under five [17, 18]. Malnutrition can manifest in six different forms, including wasting, stunting, underweight, overweight, obesity and micronutrient deficiency.

Wasting (Acute Malnutrition): Low weight for height indicates severe malnutrition and undereating or weight-loss disorders. The World Health Organization (WHO) has set up an emergency threshold of 8.4% for trash among kids younger than five.Today, Bangladesh is home to around 162.9 million people, of whom over 15.2 million are children aged between a newborn and 59 months or older [16]. Reports indicate that 9.5 million and more children suffer from stunting, 56% suffer from underweight, and over 17% experience waste [19]. Acute food shortages, particularly following natural disasters such as floods and cyclones, are the root cause of this issue. Inadequate infection treatment and restricted access to healthcare are also major causes. Infants and young children's eating habits are not optimal. As a result, one becomes more susceptible to illnesses. The death rate is greater, particularly for children under five. Weakened immune function is one of the long-term health concerns.

Stunting (chronic malnutrition): One dietary imbalance that might slow down an infant's development is called stunting [22]. Stunting refers to low height for age and is a result of long-term inadequate nutrition and repeated infections. In densely populated areas, issues with hunger and the dangers of stunting on the evolution of society and economy young children arise quickly [23]. Toddlers with stunting have trouble developing to the fullest extent possible on all fronts, including their physical and peripheral growth. Stunting may harm children under five health, education, and productivity[24]. Stunting remains a significant problem, affecting over 28% of children under five, though it has decreased over the past decade. It is more common in rural areas and among children from lower-income families. School-age children, teenagers, and even adults are affected by stunting [27]. Consequently, inadequate maternal nutrition occurs during pregnancy. Inadequate breastfeeding and complementary feeding practices. Poor sanitation, limited healthcare access, poverty, and food insecurity lead to recurring illnesses including diarrhea and respiratory illnesses. This causes the latter to have impaired physical growth and development. Delayed cognitive development can result in subpar academic performance and diminished earning potential in adulthood. Chronic illnesses are more likely to occur later in life.

Underweight children are underweight for their age. This could be a result of both stunting and malnutrition. Approximately 22% of infants under the age of five in Bangladesh are underweight. Insufficient food intake, frequent illnesses, and poor care methods all contribute to children being underweight. Being underweight can increase morbidity and mortality in children. Worldwide, 2.6 million children perish each year as a result of starvation, and 45% of those under the age of five died from undernourishment [20, 21]. It has an effect on both cognitive development and academic success. Conditions in which an individual is excessively overweight in relation to their height. Unusual fat accumulation might do damage to one's health.

Body Mass Index (BMI) is regarded as a reliable measure of nutritional status [25]. The body mass index (BMI),a weight-to-height ratio, typically classifies people as obese or overweight. It is defined as the ratio of a person's weight in kilograms to their height in meters squared (kg/m2). If the body mass index (BMI) is 25 or higher, adults are classified as overweight, and obese if their BMI is 30 or higher. BMI standards for overweight and obesity vary with age in children and adolescents. The disparity between excessive and inadequate energy consumption leads to overweight and obesity. People worldwide consume more energy-dense foods and drinks, rich in fats and carbohydrates, and engage in less physical activity. Addressing the root causes, such as inadequate nutrition, a lack of physical activity, and socioeconomic factors, is critical to reducing the increase in adolescent obesity. Comprehensive measures including parents, schools, governments, and communities are vital to fighting the growing trend of childhood obesity and ensuring a healthy future for children.

Micronutrient deficiencies, also called "hidden hunger," occur when the body does not get enough of certain vitamins and minerals. These include elements like A, zinc, iodine, and iron, all of which are important for proper growth and development despite eating enough calories. The human body also relies heavily on vitamin fulfillment. Vitamins help with the development, control, and augmentation of biological activities, while minerals have a role in many stages of energy metabolism responses, growth, and body maintenance [26]. Unlike micronutrient deficiencies (protein, fat, and carbohydrates), micronutrient deficiencies are more difficult to identify since they do not often show obvious symptoms but might have long-term health repercussions. A lot of children only eat rice, with little access to fruits, vegetables, and diets high in protein. People with iron deficiency anemia may feel tired, have weak immune systems, and have trouble thinking and remembering things. Insufficient vitamin A intake may cause vision problems and increased susceptibility to infections. Long-term zinc deficiency stunts growth and raises the risk of diarrhea and pneumonia. About 40% of children under the age of five suffer from anemia and vitamin A deficiency is especially common.

1.2. Motivation:

Child malnutrition remains a critical global issue in 2024, exacerbated by factors such as conflict, inequity, and climate crises. An approximate 181 million children under the age of five—or nearly one in every four—are experiencing acute nutritional poverty. Malnutrition, particularly wasting, may be fatal because it causes children to be dangerously underweight for their height. They are thus more likely to have general malnutrition. These children deprive themselves of essential nutrients found in fruits, vegetables, and proteins because they often only consume two kinds of foods: milk and starchy staples (Help ChildrenNow | UNICEF USA).

Sub-Saharan Africa and South Asia suffer the highest burden, affecting tens of millions of children. In conflict-affected places like Somalia and the Gaza Strip, up to 90% of children suffer from acute malnutrition (Help ChildrenNow | UNICEF USA).

In the 1980s and 1990s, the number of people who were malnourished around the world went down. This was due to improvements in health care, farming, and foreign programs run by organizations like the WHO and the World Food Program. Famines and nutritional crises, such as the one in Ethiopia in 1984, brought the issue to public notice, although they were caused by poverty, poor leadership, and warfare in places like the Horn of Africa.

2000s: In the early 2000s, the Millennium Development Goals (MDGs) served as a significant tool in the fight against hunger. There were advances in several places between 2000 and 2015. For example, when economies in East Asian countries improved, the incidence of malnutrition fell dramatically. Child malnutrition remained common in Sub-Saharan Africa and South Asia, where undernutrition (such as stunting and wasting) continued.

A variety of issues, including global conflict resolution, climate change, economic instability, and improvements to the healthcare system, will greatly influence the future of child malnutrition. We anticipate that climate change will exacerbate food insecurity, particularly in Sub-Saharan Africa, South Asia, and Central America. Rising temperatures, droughts, floods, and deserts will limit agricultural production, raising food costs and increasing starvation. Climate change has the potential to considerably raise child hunger rates in susceptible countries by 2050 (UNICEF DATA).

In Asia, particularly in South and Southeast Asia, child malnutrition remains a significant issue, despite recent progress. In this scenario, there are notable country variations in the high rates of wasting, stunting, and nutritional deficiencies.
South Asia has some of the world's highest child malnutrition rates, particularly in India, Bangladesh, Pakistan, and Afghanistan. Roughly half of all stunted children worldwide originate from this region.

Despite its economic affluence, India still suffers from child malnutrition. “About 35% of kids under five are undersized, and 19% are wasted,” says the National Family Health Survey-5 (NFHS-5). Poverty, lack of access to nutritious food, filth, and poor healthcare are among the factors (UNICEF DATA).

Bangladesh and Pakistan also face significant obstacles. Stunting affects 28% of children in Bangladesh and approaches 40% in Pakistan (UNICEF DATA).

1.3. Rationale of the Study:

Given that malnutrition is a significant issue worldwide, particularly on the Asian continent, machine learning techniques are especially well-liked for precisely predicting the risk of any illness [28]. My goal is to develop a viable solution using a machine learning algorithm, focusing solely on Asian data sets while also incorporating data sets from other parts of the world.

1.4. Research Gap:

Every day, the number of children dying increases worldwide. A saying states that "Today's child is tomorrow's future." As a result, we can now study how to anticipate child malnutrition based on global data rather than just data from Asia. Malnutrition among Bangladeshi children aged 0 to 5 is on the rise. Our next research will focus on Bangladeshi hilly districts and coastal areas, which have poor education rates and a significant risk of child marriage. Children under the age of five are more susceptible to malnutrition.

1.5. Research Question:

This research is primarily concerned with the following issues:

1. Which model is more accurate at detecting child malnutrition?

2. Which model has the lowest accuracy rate in detecting childhood malnutrition?

3. How efficient are these models at identifying problems other than child malnutrition?

4. How does malnutrition contribute to child mortality rates, and what interventions can reduce these rates?

5. How can healthcare systems be strengthened to identify and treat malnutrition in children early?

1.6. Expected output:

Experts predict that child malnutrition will have detrimental effects on both national and international development, along with serious short- and long-term health repercussions.

Addressing child malnutrition with effective treatments and policies is critical to breaking the cycle of poverty, improving health outcomes, and promoting long-term development.

1.7. Research Objective:

The goal of child malnutrition research should be to identify the issue's scope, causes, effects, and potential solutions. We aim to determine the current incidence of stunting, wasting, and underweight children in a specific area or worldwide. Determine the demographic (age, gender, socioeconomic status) and geographic (urban/rural) distribution of childhood malnutrition. The objective is to evaluate the efficacy of current treatments (dietary plans, medical campaigns, and educational projects) aimed at reducing childhood malnutrition. Our mission is to offer locally led solutions that address the root causes of malnutrition and promote long-term development. Explore the connections between child malnutrition and other global issues such as climate change, food security, and poverty. We aim to investigate how tackling malnutrition could aid in accomplishing worldwide development objectives like the United Nations Sustainable Development Goals (SDGs). These research objectives will help to focus the investigation on comprehending the complex nature of malnutrition in children using machine learning, figuring out its underlying causes, and developing efficient interventions and guidelines.

CHAPTER 2

LITERATURE REVIEW

2.1. Literature Review

A major issue facing the Asian continent is the high rate of malnutrition among children between the ages of 0 and 5. Its impact in Bangladesh and other developing nations is growing daily. They might create a gadget. They will thus develop a tool to detect youngsters who are more vulnerable and malnourished [1]. The fundamental issue is that childhood malnutrition causes many different health issues in children, and it is sometimes too late to recover from the challenges. Thus, malnutrition is one of the leading causes of illness in underdeveloped countries. Bangladesh has already supplied a sizable quantity of food to help the malnourished children in this condition. Malnutrition accounts for about half of child deaths in Asia and Sub-Saharan Africa. This study employed the secondary records from the 2014 Bangladesh Demographic and Health Survey (BDHS), which were nationally representative [2]. This study uses a single ML classification algorithm. We refer to this algorithm as K-Nearest Neighbour (K-NN). I believe that the outcome of this machine learning method is not very accurate. We employ a multitude of ML techniques, such as Random Forest, Support Vector Machine, Naive Bayes, Decision Tree, Logistic Regression, among others, to effectively achieve the desired outcome. Undernutrition is a major global public health issue that leads to high mortality rates. The nutritional status and growth of their child are the primary concerns of every parent, as the food they ingest is a factor in the brain's development and memory. When a child's food intake is insufficient to meet his nutritional requirements, he is at risk of malnutrition [3]. Every five years, the world conducts the 2016 Ethiopian Demographic and Health Survey, the most recent in its series. In an effort to better the health of Ethiopian mothers and their children, this survey gathers information on a range of demographic and health-related topics from a representative sample of residential households. The findings indicate the need for further efforts to enhance women's access to clean water, food security, and overall social and economic well-being in Ethiopia. There has to be a focus on policies and treatments for babies born with low birth weights, infants older than 30 months, and babies born to moms who are underweight [4].

One of the biggest health problems for Bangladeshi babies ages 0 to 5 is not getting enough food. A kid is considered malnourished when their caloric and protein intake is inadequate to support proper growth and development. We identified and predicted the risk variables for malnutrition (stunted, depleted, and underweight) using ML algorithms. This research used malnutrition data from the Bangladesh Demographic and Health Survey (2014) [5]. One of the biggest threats to health and wellbeing in developing nations, such as Bangladesh, is malnutrition. To improve female education standards, the Bangladesh government has instituted a variety of programs, including stipends and free education. Our research also revealed a higher risk of waste and underweight in children whose fathers only completed elementary and secondary school, compared to those whose fathers had a higher level of education [6]. Poorer consumption of essential nutrients linked to food insecurity leads to poor growth metrics and an increased risk of malnutrition. By eliminating the intricate variables that contribute to food insecurity and guaranteeing that children receive sufficient nutrients, it is essential to improve their health and prevent malnutrition. By emphasizing the need to address both macro- and micronutrient deficiencies in those experiencing food insecurity, this study contributes to the expanding body of data in favor of a holistic strategy to fight malnutrition [7]. Still one of the most important worldwide health and welfare problems is malnutrition. Trains, tests, and chooses the best ML classifier to identify stunting in Zambian children under five years old. It also finds key characteristics using data from the 2018 Zambia Demographic Health Survey (ZDHS). With a score of 79% in testing and 61.6% in training data, RF was the algorithm that demonstrated the highest level of accuracy. Significant predictors of stunting in children under five years old were also found, but this time they related to the socioeconomic position of the mothers and the children [8]. This study examines the various socioeconomic factors that impact children's nutritional health. Therefore, data mining to cluster risk factors for under-five malnutrition is our primary study emphasis. We examine BDHS 2014 data using random forest and decision tree models. The RF accurately predicts stunting and underweight at 70.1% and 72.4%. To nurture a healthy future generation, we can apply many target-based and fact-finding treatments. Today's youngsters may become tomorrow's leaders [9]. Childhood obesity, heart disease, type 1 diabetes, and osteoporosis are just a few of the long-term health issues that many young people suffer from because of their unhealthy eating habits. We can increase the likelihood that children will live long, healthy lives free of disease if we teach them the importance of eating all the food groups. When it comes to increasing nutritional status, this study will be a huge help to both individuals and healthcare providers. Government health services will aid in improving people's diets [10]. Every year, 12 million children under the age of five pass away; the majority of these children reside in underdeveloped nations. The meticulous Nigeria Multiple Indicator Cluster Survey (MICS6) 2021 provided data for the study. This survey uses four machine learning algorithms, with the KNN model identified as the most successful predictor. This research represents a significant breakthrough in our knowledge of and ability to address childhood malnutrition in Nigeria. The statement underscores the significance of merging cutting-edge analytical methods with public health investigations to tackle intricate societal problems [11]. The prompt detection of childhood malnutrition is crucial for the prevention of many instances and the provision of suitable treatment to hinder the progression of the disease. The Sawah Besar Community Health Service provided 360 patients with 4 traits' worth of medical records. This research investigates three supervised machine learning approaches. When it comes to accuracy, Decision Tree is on top with an impressive 89.87% [12]. Every country's government places a premium on children's health in order to build a strong nation. "A STRONG MIND IN A STRONG BODY" is a saying that suggests that strong bodies breed powerful civilisations. We should detect malnutrition in children as soon as possible. As an extra benefit, when medical treatments are widely available and cheaply priced, it is simple to get expert advice on how to treat various types of malnutrition. Technology can help parents find out right away if their child isn't getting enough food. Through the use of a machine learning model, someone who is not a medical professional can also help by using a machine learning model. An effective and precise machine learning model is required [13]. Balance is necessary for healthy eating. Mothers and children in Bangladesh are malnourished. Undernutrition and obesity are examples of malnutrition among women. According to the World Health Organisation (WHO), 190 crore are overweight and 462 billion are underweight. Despite being a serious problem, Bangladeshi malnutrition is decreasing [14]. Kids' growth and development are crucial from 1 to 60 months. Childhood malnutrition detection is crucial for lifetime health. Thus, creating a model for early identification and prediction of childhood malnutrition is the goal here. This study used a random sample of 574 children under the age of five from the Lunugala area to construct a predictive model for malnutrition using ensemble learning techniques. This study uses nine ML models, and the high accuracy rate is 93%. Ensemble learning is superior at predicting nutritional status, according to the study [15].

CHAPTER 3

RESEARCH METHODOLOGY

Child malnutrition is a condition in which a child's diet lacks the critical nutrients required for normal growth and development. It may result in undernutrition (wasting, stunting, or being underweight), micronutrient deficiencies (for example, iron or vitamin A deficiency), or overnutrition. Poor eating habits, illnesses, insufficient healthcare, and poverty are also contributing factors. Malnutrition leads to decreased physical and cognitive development, a compromised immune system, as well as an elevated risk of both disease and death. Breastfeeding, a balanced diet, access to healthcare, and nutritional education are examples of preventive strategies. This section provides a full overview of the dataset, data preparation, and recommended approaches for preprocessed data using different languages and algorithms.

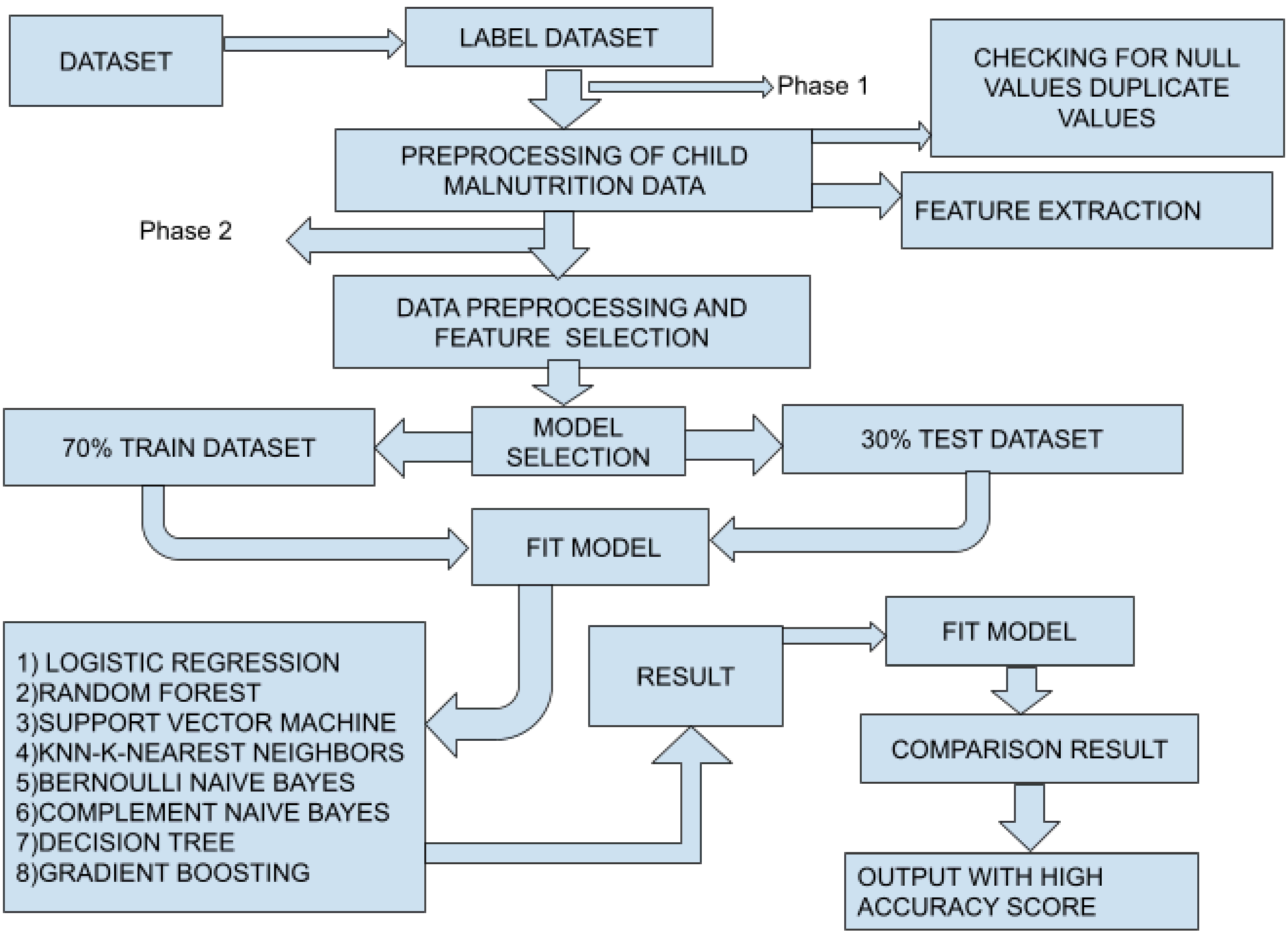

Figure 3.1: Workflow of this study

3.1. DATA COLLECTION:

The comprehensive datasets were gathered. After that, they looked through each dataset to find an appropriate one. In the end, they were able to get raw data from public sources. We aim to encompass all continents and subcontinents where undernourished children reside. We have compiled a diverse array of demographic data, including birth records, income scales, height, and weight records, from diverse regions throughout Asia. Our model performs an analysis of 837 data points and 12 features from the dataset. In this dataset, there are already entities such as overweight, underweight, and stunning. Thus, anthropometric analysis was omitted.

3.2. DATASET DESCRIPTION:

This child starvation collection gathers data from various areas, genders, and age groups to evaluate children's overall health. It examines nutritional trends such as underweight, stunted, wasting, and overweight conditions by combining health, population, and social data. Policymakers, academics, and global health organizations may use this information to track the prevalence of child malnutrition and create focused interventions.

Features of the Dataset:

1. ISO Code:

Description: The ISO 3166-1 country code is a standardized two- or three-letter identifier that identifies the child's home country.

Purpose: To relate each child's data to their nation, allowing for country-level research and comparisons.

2. Sex:

Description: The letters "M" for male and "F" for female indicate the child's biological sex.

Purpose: This study aims to explore gender differences in malnutrition rates, given that biological or sociocultural factors may impact boys and girls differently.

3. Age:

Description: The child's age, usually expressed in months or years.

Purpose: Nutritional demands and development rates vary by age; hence, this feature helps identify age-specific malnutrition trends (e.g., increased susceptibility in children under five years)

4. Height:

Description: The kid's height or length (depending on whether the youngster is standing or lying down), commonly measured in cm.

Purpose: The purpose is to calculate height-for-age (stunting) and height-for-weight (wasting), which are important markers of chronic and acute malnutrition.

5. Weight:

Description: We state the child's weight in kilograms.

Purpose: This important anthropometric test aims to detect various types of malnutrition by evaluating weight-for-height (wasting) and weight-for-age (underweight).

6. Continent:

Description: The description includes the child's continent of residence, be it Africa, Asia, or South America.

Purpose: We do this to examine hunger trends in a broader region and compare them across countries.

7. Global Bank:

Description: This most likely refers to the World Bank's classification of the nation's income, which includes poor, lower, medium, and high incomes.

Purpose: Correlates malnutrition trends with the child's nation's economic position, offering insights into the impact of economic circumstances on nutrition.

8. Wasting:

Description: A binary or categorical indicator that indicates if the youngster is wasting (low weight for height), often represented as "yes/no" or "1/0."

Purpose: The goal is to identify children who are severely undernourished and at risk of rapidly deteriorating their health.

9. Overweight

Description: A binary or categorical indicator that indicates whether the child is overweight based on their body mass index (BMI) or weight for height.

Purpose: Tracks the growing issue of juvenile obesity, which may coexist with undernutrition in certain areas.

10. Sunting:

Description: An indication, either binary or categorical, indicates whether the child is stunted (low height for age).

Purpose: The goal is to define chronic malnutrition, a condition often resulting from prolonged deprivation or undernutrition during early life.

11.Underweight:

Description: An indicator, either binary or categorical, determines whether a child is underweight (low weight for age).

Purpose: Determines whether the child's overall weight is age-appropriate and points out potential undernutrition in general.

12. Status:

Description: An overall indication of the child's nutritional state or condition. This might label youngsters as "normal," "malnourished," "at risk," and so on.

Purpose: Provides a summary assessment of the child's health by combining the numerous anthropometric parameters to determine overall nutritional status.

This dataset provides a complete platform for analyzing child malnutrition at a global or regional level.

	ISO code	Sex	Age	Height	Weight	Continent	World Bank	Wasting	Overweight	Stunting	Underweight	Status
0	AFG	1	2	93	16	Asia	Low Income	18.2	6.5	53.2	44.9	Stunting
1	AFG	1	4	97	15	Asia	Low Income	8.6	4.6	59.3	32.9	Stunting
2	AFG	0	4	90	12	Asia	Low Income	9.5	5.4	40.9	25.0	Stunting
3	ALB	0	2	78	16	Europe	Upper Middle Income	8.1	9.5	20.4	7.1	Stunting
4	ALB	0	4	99	13	Europe	Upper Middle Income	12.2	30.0	39.2	17.0	Stunting
5	ALB	1	2	71	13	Europe	Upper Middle Income	7.3	25.2	27.0	6.6	Stunting
6	ALB	0	1	94	11	Europe	Upper Middle Income	9.4	23.4	23.1	6.3	Overweight
7	DZA	1	2	92	10	Africa	Upper Middle Income	4.0	NaN	16.9	8.0	Stunting
8	DZA	0	4	101	15	Africa	Upper Middle Income	7.1	8.7	22.9	9.2	Stunting
9	DZA	0	1	86	17	Africa	Upper Middle Income	9.6	13.2	22.5	11.3	Stunting

Figure 3.2: Data table

3.3. DATA PREPROCESSING:

3.3.1 DATA HANDLING:

Data handling for child malnutrition includes preparing the dataset to ensure the analysis's accuracy and reliability. This procedure involves dealing with missing data by either imputing , as well as finding outliers in critical variables such as height for age or weight for height, which might suggest possible inaccuracies. We find and eliminate duplicate items to avoid biased outcomes. Standardization corrects errors such as mismatched ages, nutritional indicators, and units of measurement. We also examine the homogeneity of categorical data, such as gender. Addressing these challenges improves the quality of malnutrition data, resulting in more accurate insights and policy recommendations. During phase 1, the dataset was not entirely prepared for the fitting and training of our model, as it contained some null and duplicate values. So, initially, we utilized the most popular Python library, "NumPy," to handle the missing and duplicate data.

Additionally, they restricted our global dataset to Asia. We have completed the feature extraction procedure. They just constructed a fake dataset of "World Bank" attributes (low income, upper middle income, lower middle income) in phase 2. They then integrated this new dataset with the existing one. We then removed the "World Bank" attribute. We then converted our dataset into a csv file with the correct attribute structure.

3.3.2 FEATURE SELECTION:

In the context of child malnutrition, feature selection entails determining the most important variables from a dataset in order to increase prediction model accuracy and efficiency. Age, weight, height, and gender are common characteristics for measuring malnutrition indicators such as stunting (height for age), wasting (weight for height), overweight, and underweight. Socioeconomic position, maternal education, dietary variety, availability to healthcare, and geographic location are all important elements to consider when considering malnutrition risk factors. Data analysis becomes more targeted, lowering noise, enhancing model performance, and offering more precise insights into the main causes of child malnutrition by limiting the number of characteristics to those that are most relevant. During feature selection in data preprocessing, the univariate selection method is used, where the function is f-classif and the algorithm is SelectKBest.

3.3.3 SPLIT THE DATASET:

Splitting the dataset in child malnutrition analysis is a critical step to ensure the reliability of predictive models. We typically separate the dataset into two subsets: the training set and the testing set. The training set generally contains 70-80% of the data, trains the model to identify patterns and correlations among variables such as age, weight, and indications of malnutrition. Finally, we utilize the test set (20-30%) to analyze the model's performance and offer an impartial estimate of anticipated accuracy. This method ensures the reliability and successful application of the child malnutrition detection model with fresh, previously unknown data.

3.4 MODEL SELECTION:

They investigated numerous acceptable machine learning algorithms for our dataset that might aid in the development of a strong ML model and the prediction of high accuracy rates. They chose eight ML algorithms, including Random Forest, Support Vector Machine, Logistic Regression, K-nearest Neighbors, Bernoulli Naive Bayes, Complement Naive Bayes, Decision Tree, and Gradient Boosting, after carefully weighing the benefits and drawbacks of each approach. The input was then used to train our model and validate accuracy (the percentage of correct forecasts to all predictions). After completing the comparison process, we could identify the optimal method.

3.4.1.RANDOM FOREST (RF) CLASSIFIER:

In addition to being very stable, Random Forest is not a biased algorithm. Random Forest performs exceptionally well when faced with both absolute and mathematical conditions, incomplete or poorly scaled information, and large datasets. We get excellent accuracy via Random Forest. Random Forest takes a significant amount of time to train our model. We use it to cultivate a large number of trees and then harvest their fruits.

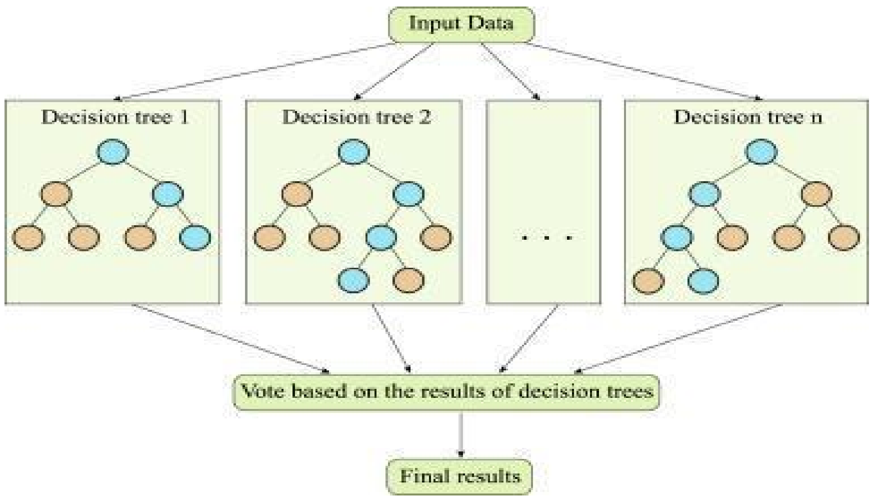

Figure 3.3 : Random Forest classifier working process

How RF Works:

- **Building Multiple Decision Trees**:
  - Random Forest generates a huge number of decision trees (typically hundreds or thousands) during the training phase.

- Each tree receives training from a random subset of the data obtained through bootstrapping (sampling with replacement).
- For each tree, we choose a random group of traits at each split (node). This introduces randomization and ensures the trees remain uncoupled, thereby preventing overfitting.

  For instance, if there are ten features, the system may randomly select three characteristics from each node to determine the division.

- **Making Predictions**:
  - After the tree forest is built, the RF makes predictions by aggregating the outputs of each individual tree.
  - For classification problems, each decision tree gives a "vote" to a class. A majority vote determines the final prediction, selecting the class with the most votes across all trees.
  - In regression tasks, we average the predictions from each tree to arrive at the final prediction.

Mathematical Operations:

- **Data Representation**:
  - X = feature matrix.
  - Y = target labels.
- **Bootstrap Sampling**:
  - For each tree, randomly sample the dataset with replacement to form subset

$$D_b \subset D .$$

- **Random Feature Selection**:
  - At each node, randomly select mmm features from the total ppp features for splitting (e.g., m≪p).

- **Node Splitting (Gini Impurity/Entropy)**:
  - Calculate Gini impurity for classification:

$$Gini(t) = 1 - \sum_{i=0}^{1} p_i^2$$

- Where p_i is the proportion of malnourished (1) or not malnourished (0) children at the node.
- Or use Entropy:

$$Entropy(p) = -\sum_{i=0}^{1} p_i log_2(p_i)$$

- Select the split that minimizes impurity.

❖ **Tree Construction**:

- Recursively split nodes until stopping conditions (e.g., minimum samples, max depth) are met.

❖ **Prediction**:

- For a new child's feature vector xxx, each tree provides a prediction $h_i(x)$, and the final output is by majority voting:

$$H(x) = mode(h_1(x), h_2(x), \ldots\ldots\ldots, h_T(x))$$

❖ **Out-of-Bag (OOB) Error**:

- Use OOB samples to estimate model accuracy.
- OOB samples are used to estimate error, calculating the proportion of misclassified samples:

$$OOB\ Error = \frac{1}{n}\sum_{i=1}^{n} \Pi(H_{OOB}(x_i) \neq y_i)$$

Where $H_{OOB}(x_i)$ is the OOB prediction and y_i is the true label.

3.4.2. SUPPORT VECTOR MACHINE (SVM) CLASSIFIER :

Support Vector Machine is a highly effective technique that can clearly separate data and classes, while also offering versatility in selecting related functions. It is capable of handling large sizes. By using SVM in areas with a lot of dimensions, it is possible to classify data very well. performs poorly on huge datasets and is unable to function properly in situations when classes are not distinct.

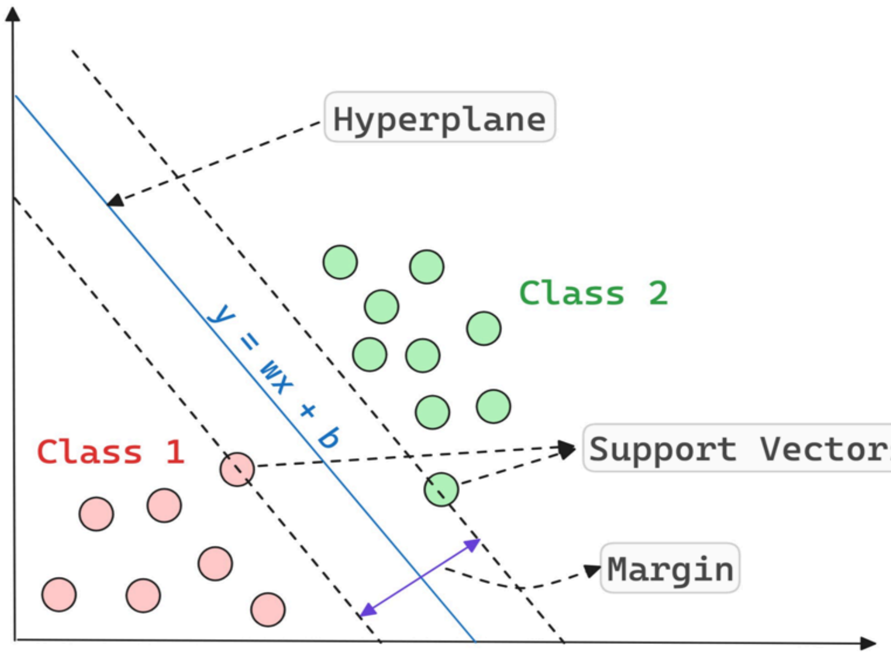

Figure 3.4: Support Vector Classifier Working Process.

How SVM Works:

- **Hyperplane:** Finding the best hyperplane between two classes is the goal of SVC. The margin, or the separation between the hyperplane and the nearest data points in each class (sometimes referred to as support vectors), is maximized by the optimal hyperplane.
- **Maximizing Margin:** Margin maximization strengthens the classifier's resilience to new data points. The classifier's generalization improves as the margin grows.
- **Soft Margin:** If the data is not completely separable, SVC allows for some misclassifications by leveraging slack variables to create a "soft margin" that balances favorable classification with margin size.
- **Kernel approach:** To transform non-linearly separable data into a higher-dimensional space suitable for linear separation, SVC employs the kernel technique. Polynomial and RBF (Radial Basis Function) are two typical kernel types.

Mathematical Summary:

- **Decision Function:** The classifier's decision function is:

$$f(x) = sign(\omega \cdot x + b)$$

where the sign function determines the class label (malnourished or not).

❖ **Maximizing Margin:** The goal is to maximize the margin by reducing:

$$\frac{1}{2}||\omega||^2$$

subject to:

$$y_i(\omega \cdot x_i + b) \geq 1$$

❖ **For Soft Margin** (with slack variables):

$$min \frac{1}{2}||\omega||^2 + C \sum_{i=1}^{n} \xi_i$$

with $\xi_i \geq 0$

❖ **Kernel Trick**: For non-linear data, the decision boundary becomes:

$$f(x) = sign\,(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b)$$

where $K(x_i, x)$is the kernel function (e.g., RBF or polynomial).

3.4.3. K-NEAREST NEIGHBORS (K-NN) CLASSIFIER :

The K neighbors algorithm is very easy to use since it only needs two inputs. Consequently, it is simple to implement. It is possible for K-NN to learn nonlinear decision boundaries. As a result, it may employ nonlinear functions to partition the data. The process of training the data is swift, enabling it to operate efficiently. It doesn't work very well. If our dataset contains a large number of dimensions, it may not function correctly. It has trouble with categorical traits.

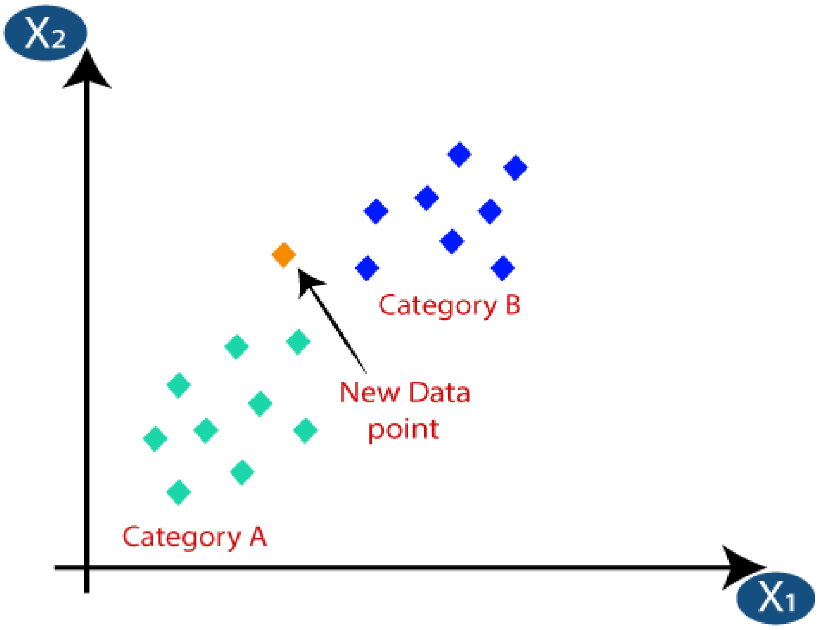

Figure 3.5: K-Nearest Neighbors Classifier Working Process.

How K-NN Works:

The K-Nearest Neighbors (K-NN) classifier is a straightforward non-parametric method for solving regression and classification issues. It compares a new data point to the majority class of its K closest neighbors in the feature space before classifying it based on previously recorded training points. Typically, K-NN uses the euclidean distance between each training point and each new point to determine its classification. Following a majority vote, it gives the K nearest points to the class with the most common neighbors. Though simple to implement, K-NN's insistence on retaining all training data and performing distance calculations at prediction time makes it potentially computationally costly for big datasets.

Mathematical Summary:

➢ **Data Representation**:
- X = feature matrix.
- Y = target labels.

➢ **Distance Calculation:**

For each new test point xxx (e.g., a new child's data), K-NN calculates the distance to all training points. A common choice is **Euclidean distance**:

$$d(x\,,\,x_i)\;=\;\sqrt{\sum_{j=1}^{d}\left(x_j - x_{ij}\right)^2}$$

Where:

- $x\;=\;(x_1\,,\,x_2,.....,x_d)$is the array of characteristics for the newborn,
- $x_i = \;(x_{i1}\,,\,x_{i2},.....,x_{id})$is the feature vector for the iii-th training child,
- d is the number of features (e.g., age, weight, height).

➢ **Finding the K Nearest Neighbors:**

Once distances between the test child x and all training children $x_1\,,\,x_2\,,......,\,x_n$ are computed, K-NN selects the KKK nearest neighbors by sorting the distances and picking the closest KKK points.

Let:

- $N_k(x)$be the set of the K nearest neighbors for x**.**

➢ **Majority Voting for Classification:**

To classify the new child, K-NN uses majority voting. The class label for the test child x is determined by the most common class among its K-NN..

The predicted label $\hat{y}$ is given by:

$$\hat{y}\;=\;argmax_{y\epsilon\{-1\,,\,+1\}}\;\sum_{x_i\epsilon N_k(x)} 1(y_i\;=\;y)$$

Where:

- $1(y_i\;=\;y)$is an indicator function that equals 1 if the neighbor x_i has label y, and 0 otherwise.
- $N_k(x)$is the set of the K-nearest neighbors.

➢ **Weighted Voting (Optional):**

In some cases, K-NN can use weighted voting, where closer neighbors have a higher influence on the decision than distant neighbors. A common weighting function is based on the inverse of the distance:

$$\hat{y} = argmax_{y\epsilon\{-1,+1\}} \sum_{x_i \epsilon N_k(x)} \frac{1(y_i = y)}{d(x, x_i)}$$

This gives more importance to neighbors that are closer to the test point.

3.4.4. LOGISTIC REGRESSION (LR) CLASSIFIER :

The logistic regression method is highly explicable, producing outputs with proper calibration and expected probability. The model's training and prediction components are extremely rapid; scaling is unnecessary given the dimensions. With a limited number of observations, it can function quite effectively. It guesses what the binary outcome of a certain independent variable will be. The LR uses a large-dimension dataset to overfit the model, but the outcome is inaccurate.

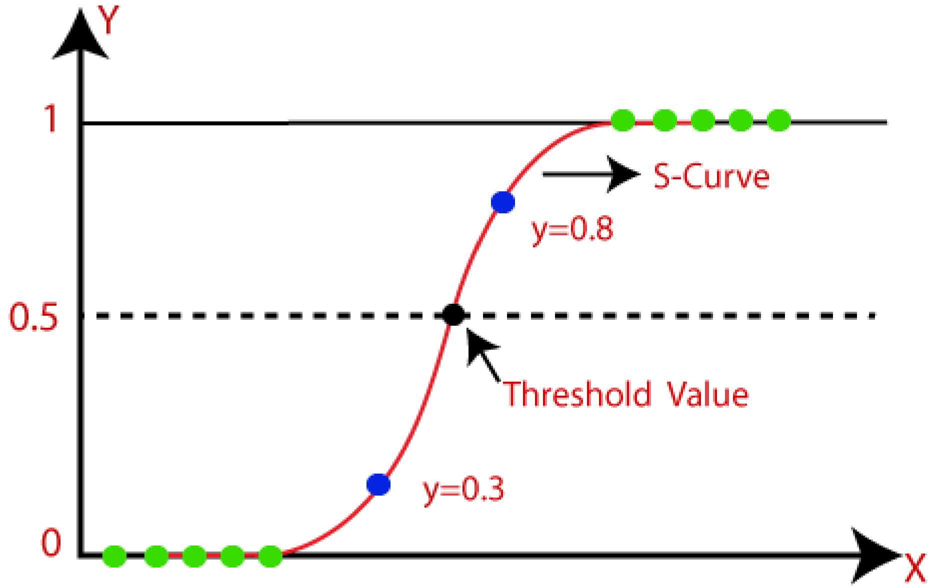

Figure 3.6: Logistic Regression Classifier Working Process.

How LR Works:

Used for binary classification tasks, logistic regression is a supervised learning algorithm. It forecasts the likelihood of a binary result (e.g., malnourished or not) based on input variables such as age, weight, and height. The model operates by generating a weighted sum of the input characteristics and applying a sigmoid function to this total, yielding a probability between 0 and 1. The model predicts the positive class (malnourished, for example) if the likelihood is higher than a threshold, usually 0.5; if not, it predicts the negative class. In classification difficulties, logistic regression is popular because it is simple to understand and use.

Mathematical Summary:

- **Feature Representation:**

  let:

  ★ $X = \{x_1, x_2, ..., x_n\}$ represent the feature.

  ★ $Y = \{y_1, y_2, ..., y_n\}$ represent the target labels.

- **Linear Combination of Features**:

  Logistic regression computes a linear combination of the input features:

  $$z = w_1 \cdot x_1 + w_2 \cdot x_2 + ... + w_d \cdot x_d + b$$

  Where:

  ★ $w_1, w_2, ..., w_d$ are the weights (or coefficients) for the features,

  ★ $x_1, x_2, ..., x_d$ are the feature values (e.g., height, weight, etc.),

  ★ b is the bias or intercept term.

  This linear combination z serves as the input to the next step, which is applying the sigmoid function.

- **Sigmoid Function (Logistic Function)**:

  The logistic function (sigmoid function) maps the linear combination z to a probability value between 0 and 1. The sigmoid function is given by:

  $$p(y = 1|x) = \frac{1}{1+e^{-z}}$$

  Where:

  ★ $z = w_1 \cdot x_1 + w_2 \cdot x_2 + ... + w_d \cdot x_d + b,$

  ★ $p(y = 1|x)$ represents the probability that the child is malnourished.

- **Decision Boundary**:

Logistic regression uses a threshold to classify a child as malnourished or not. Typically, a threshold of 0.5 is used:

$$\hat{y} = \begin{cases} 1 \; if \; p(y=1|x) \geq 0.5 \\ 0 \; if \; p(y=1|x) < 0.5 \end{cases}$$

Where:

★ $\hat{y} = 1$ means the child is predicted to be malnourished.

★ $\hat{y} = 0$ means the child is predicted to be not malnourished.

- **Cost Function**:

To train the logistic regression model, a cost function (log-loss or binary cross-entropy) is minimized. The cost function for a single training example is:

$$Cost\,(h(x), y) \; = \; -\,[ylog(h(x)) \; + \; (1 - y)log(1 - h(x))]$$

Where:

★ $h(x) \; = \; P(y \; = \; 1|x)$ is the predicted probability,

★ y is the true label (1 for malnourished, 0 for not malnourished).

The overall cost function for the dataset is the average cost over all training examples:

$$J\,(w) \; = \; \frac{1}{n}\sum_{i=1}^{n}[-\,y_i log(h(x_i)) \; + \; (1 - y_i)log(1 \, - \, h(x_i))]$$

Where:

★ n is the number of training examples,

★ $h(x_i)$ is the predicted probability for the i-th example.

3.4.5. BERNOULLI NAIVE BAYES (BNB) CLASSIFIER :

The Bernoulli Naive Bayes (BNB) algorithm can solve binary classification questions when the input traits are either true or false. The algorithm derives from Bayes' Theorem and assumes that the features are independent of the class. You can use the Bernoulli Naive Bayes algorithm to categorize malnourished infants based on binary variables like the presence or absence of specific risk markers (e.g., stunted growth, insufficient food).

How BNB Works:

1. Based on input characteristics, the Bayes' Theorem determines the probability of a class. The classifier assumes that the characteristics are conditionally independent, which means that the presence or absence of one does not impact the other features.
2. For each class (e.g., malnourished or not malnourished), the model computes the conditional probability of seeing each characteristic (binary) and then utilizes the product of these probabilities to derive the likelihood that the data belongs to that class.
3. The forecast is based on the class with the highest posterior probability.

Mathematical Summary:

- **Bayes' Theorem:**

  Bayes' Theorem: The classifier computes the posterior probability of a class $y \epsilon \{0, 1\}$ (malnourished = 1, not malnourished = 0), given binary features $X = \{x_1, x_2, ..., x_d\}$:

$$P(y|X) = \frac{P(X|y)\, p(y)}{P(X)}$$

  Where:

  - $P(y|X)$ is the posterior probability of the child being malnourished given the features,
  - $P(X|y)$ is the likelihood of the features given the class,
  - $p(y)$ is the prior probability of the class,
  - $P(X)$ is the evidence, which is the total probability of observing the features.

- **Likelihood of Features**:

  BNB assumes features are conditionally independent, therefore likelihood is the product of feature probabilities:

$$P(X|y) = \prod_{i=1}^{d} P(x_i|y)$$

  For binary features x_i, the likelihood is computed using:

$$P(x_i|y) = \theta_{iy}^{x_i}(1 - \theta_{iy})^{1-x_i}$$

Where:

- $\theta_{iy} = P(x_i = 1|y)$is the probability of the feature i being present given class y.

- **Posterior Probability**:

  The posterior probability for each class is calculated by multiplying the prior and likelihood:

$$P(y = 1|X) = P(y = 1)\prod_{i=1}^{d} \theta_{i1}^{x_i}(1 - \theta_{i1})^{1-x_i}$$

$$P(y = 0|X) = P(y = 0)\prod_{i=1}^{d} \theta_{i0}^{x_i}(1 - \theta_{i0})^{1-x_i}$$

  Where:

  - $P(y = 1)$ and $P(y = 0)$ are the prior probabilities of being malnourished and not malnourished, respectively.

- **Prediction**:

  The predicted class $\hat{y}$ is the class with the higher posterior probability:

$$\hat{y} = arg\,max_{y \in \{0,1\}}\ P(y|X)$$

3.4.6. COMPLEMENT NAIVE BAYES (CNB) CLASSIFIER :

The complement Naive Bayes (CNB) classifier is a Naive Bayes version, designed to address class imbalance in classification applications. When one class is much more abundant than others in a dataset, it performs very well. Unlike standard Naive Bayes, which computes probability for each class based on feature presence, Complement Naive Bayes computes probabilities based on the class's complement, i.e., combining information from all other classes to improve target class accuracy. When predicting child malnutrition, the text classification technique may alter if the goal class (malnourished) is less common than the healthy class.

How CNB Works:

- **Complement Class:** For each class, CNB computes the feature probabilities using its complement (opposite). This helps address class inequalities by weighing the traits.
- **Log-Likelihood:** Using the feature probabilities from the complement class, CNB calculates the log-likelihood ratios for every class.
- **Prediction:** We expect the class with the greatest log-likelihood score to be the result. This enhances CNB's adaptability to jobs where the target class, such as malnourished children, has a low representation.

Mathematical Summary:

- **Feature Representation**:
  - $X = \{x_1, x_2, ..., x_n\}$ be the feature vector for a child,
- **Complementary Probabilities**:

  Classifiers use CNB to predict class complements.

$$P(x_i|y') = \frac{\sum_{y'=y} x_i \cdot N(y')}{\sum_{y'=y} N(y')}$$

  Where:

  - $N(y')$ is the number of instances from the complement class,
  - x_i is the feature count for the complement class.
- **Log-Likelihood Ratio**:

  The classifier uses the log-likelihood ratio for each class to compute the overall score for a given input:

$$L(y|X) = log\, P(y) + \sum_{i=1}^{n} x_i log\, P(x_i|y')$$

  where

  - $L(y|X)$ is the log-likelihood for class y,
- **Prediction**:

  Model predicts log-likelihood-maximizing class $\hat{y}$:

$$\hat{y} = arg\ max_{y \epsilon \{0,1\}}\ L(y|X)$$

3.4.7. DECISION TREE CLASSIFIER:

A Decision Tree Classifier is a supervised machine learning classification algorithm. Recursively splitting the data into subsets based on feature values creates a tree-like structure with core nodes representing feature-based decisions and leaf nodes representing class labels[29]. Finding the optimum splits to categorize the data is the aim.

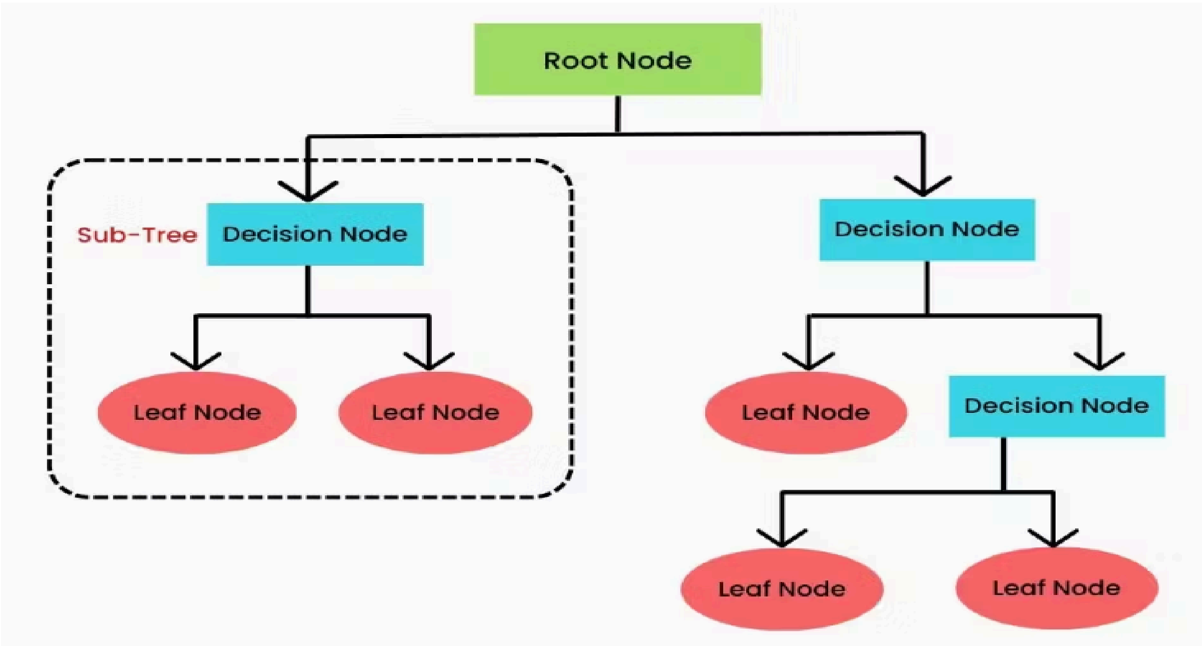

Figure 3.7: Decision Tree Classifier Working Process.

How Decision Tree Works:

- **Root Node:** The algorithm begins at the root node, where it assesses all attributes and picks the one that best divides the data into various groups (for example, malnourished vs. not malnourished).
- **Recursive Splitting:** We divide the dataset based on feature values and repeat the procedure recursively, building branches and nodes until we reach the stopping requirements (e.g., maximum depth or pure nodes).
- **Recursive Splitting:** We divide the dataset based on feature values and repeat the procedure recursively, building branches and nodes until we reach the stopping requirements (e.g., maximum depth or pure nodes).
- **Class Prediction:** The system adds new data points as choices progress from the root to the leaf node, predicting the class based on feature values.

Mathematical Summary:

- **Gini Impurity (or Entropy) for Splitting:**

  At each node, the classifier chooses the feature that minimizes the Gini impurity (or maximizes information gain in the case of entropy) [29]. For Gini impurity, the formula is:

$$Gini = 1 - \sum_{i=1}^{c} p(i)^2$$

  Where:

  - P(i) is the probability of class iii (e.g., malnourished or not malnourished) in a node,
  - c is the number of classes.

  For entropy, the formula is:

$$Entropy = - \sum_{i=1}^{c} P(i) log_2 p(i)$$

- **Information Gain**: The algorithm selects the feature that gives the highest information gain (difference in entropy before and after the split):

$$Information\ Gain = Entropy(before) - Entropy(after)$$

  This helps the decision tree find the best feature to split the data, such as child malnutrition indicators like access to healthcare or food.

- **Recursive Splitting**: The process of calculating the Gini impurity or entropy is repeated for each node as the tree splits, continuing until a stopping criterion is reached (e.g., max depth, no more splits).
- **Class Prediction**: After building the tree, a new data point follows the choices at each node based on feature values, leading to a leaf node with the majority class determining malnutrition status.

3.4.8. GRADIENT BOOSTING (GB) CLASSIFIER:

A gradient boosting classifier is an ensemble learning method that builds a series of decision trees to address the errors of the previous tree. In contrast to other ensemble approaches such as Random Forest, gradient boosting is primarily concerned with maximizing performance by gradually enhancing the predictions made by the model. It is commonly used for both classification and regression tasks.

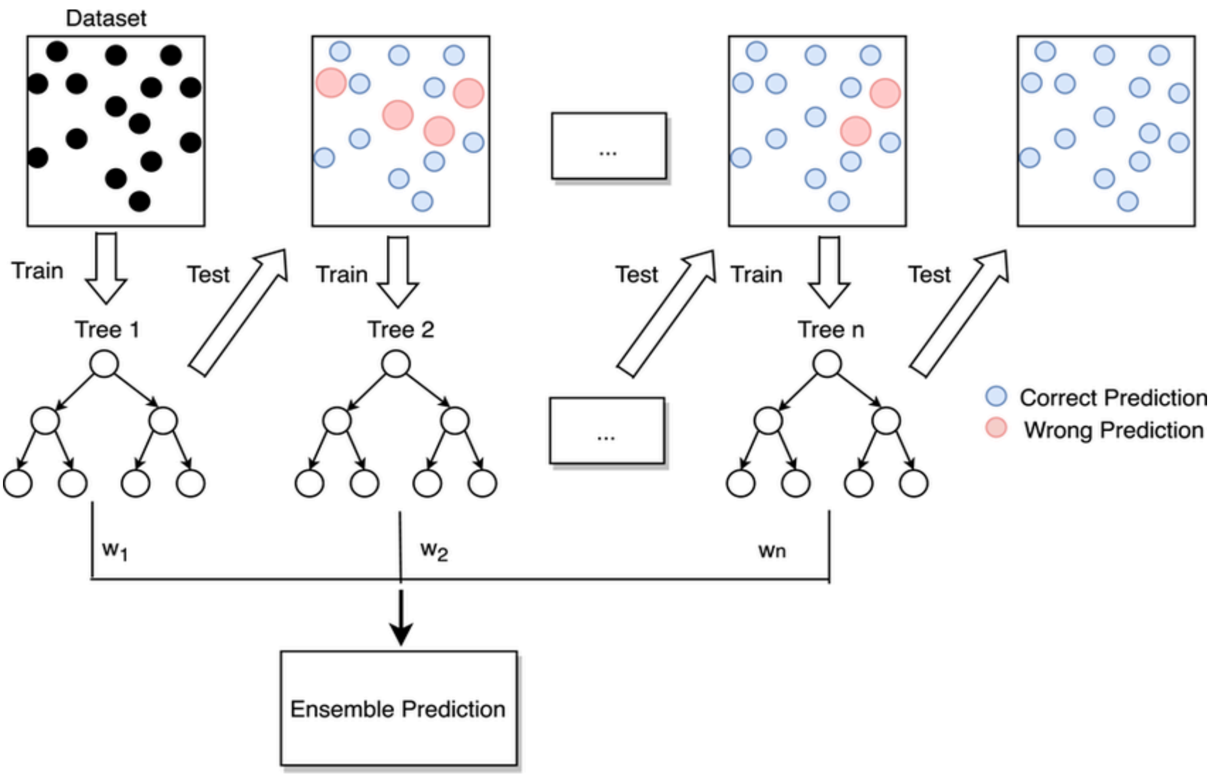

Figure 3.8: Gradient Boosting Classifier Working Process.

How GB Works:

- ❖ **Initial Prediction:** The algorithm begins with an initial prediction, which is often the mean or a basic model.
- ❖ **Residual Calculation:** For each occurrence, we calculate the residual, or error, between the actual and anticipated values.
- ❖ **New Tree Fit:** To address the shortcomings of the previous model, we set up a new decision tree based on the leftovers.
- ❖ **Update Prediction:** The output of the new tree, usually scaled by a learning rate, may be used to update the predictions.

❖ **Repeat Process:** Each tree in this process improves the model's predictions over a predetermined number of iterations.

Mathematical Summary:

- **Loss Function**:

The key idea of gradient boosting is to minimize a loss function. For classification, the most common loss function is the log loss:

$$Log\ Loss\ =\ -\sum_{i=1}^{n}[y_i\,log(p_i)\ +\ (1-y_i)log(1-p_i)]$$

Where:

  - y_i is the actual class label (1 = malnourished, 0 = not malnourished),
  - p_i is the predicted probability of class 1 (malnourished).

- **Gradient Descent**:

Gradient descent minimizes loss function.The negative gradient of the loss function relative to the prediction is calculated, which represents the direction in which the model should be improved.

For each iteration m:

$$r_i^{(m)} = -\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}$$

Where:

  - $r_i^{(m)}$ is the residual (error) at iteration m,
  - $f(x_i)$is the current prediction.

- **Updating Prediction**:

The prediction is updated by adding a fraction (controlled by the learning rate η) of the new tree's predictions to the current prediction:

$$f_{m+1}(x) = f_m(x) + \eta \cdot T_m(x)$$

Where:

- $T_m(x)$ is the new tree fitted to the residuals,
- The learning rate, denoted by η, usually ranges from 0 to 1.

- **Final Prediction:**

The final prediction for whether a child is malnourished is the sum of all trees:

$$\hat{y} = \sum_{m=1}^{M} \eta \cdot T_m(x)$$

Where M is the total number of trees, and the final output is used for classification (malnourished or not).

3.5. EVALUATION MODEL:

Evaluating a model for predicting child malnutrition entails comparing its performance against key indicators and validation approaches. This assures that the model is both accurate and generalizable to previously unknown data, which is important given that predictions in this context might have a real-world influence on health treatments.

3.5.1. Key Evaluation Steps:

- **Data Preparation:**
    1. The dataset has 70% and 30% training and testing sets. The training set develops the model, whereas the test set evaluates it.
    2. Cross-Validation: K-fold cross-validation occasionally ensures that the model's performance is independent of data division.
- **Evaluation Metrics:**

  Several performance indicators are utilized to assess a model for child malnutrition prediction, especially since it is a binary classification issue (e.g., malnourished vs. not malnourished)

    1. **Accuracy:** The proportion of right forecasts.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

When the categorization result is taken into consideration, the terms "True Positives," "True Negatives," "False Positives," and "False Negatives" are all used.

2. **Precision:** The percentage of actual positive results among all positive predictions [30].

$$Precision = \frac{TP}{TP+FP}$$

If the goal is to minimize false positives, accuracy is crucial (e.g., wrongly diagnosing a youngster as underweight).

3. **Recall (Sensitivity)**: The proportion of true positives detected by the model.

$$Recall = \frac{TP}{TP+FN}$$

High recall is important if it's crucial to identify as many malnourished children as possible [30] [31].

4. **F1 Score**: The harmonic mean of accuracy and recall, which balances the two measurements[31] [32].

$$F1\,Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

When working with an unbalanced dataset—where instances of malnutrition may be significantly less common—this is very helpful.

- **Confusion Matrix:**

The confusion matrix illustrates the distribution of predictions, thereby offering a comprehensive understanding of the model's performance. It consists of:

1. True Positives (TP): The prediction of malnourished children was accurate.
2. True Negatives (TN): The prediction was accurate, indicating that there were no malnourished children.
3. False Positives (FP) are children for whom malnutrition was anticipated but not actually observed.

4. False Negatives (FN): People often assume that children who are undernourished are well fed.

This assists in identifying the areas in which the model is making errors.

- **Dealing with Imbalanced Data:**

Because malnutrition cases are often less common in real-world data, the dataset may be lopsided. To deal with this:

1. Class Weighting: The method may be modified to provide more weight to the minority class (malnourished children).
2. Oversampling/Undersampling: SMOTE (Synthetic Minority Oversampling Technique) may be used to balance a dataset.

- **Model Selection:**
Evaluation metrics determine the optimal model. Random Forest, Support Vector Machine, KNearest Neighbors, Logistic Regression Bernolli Naive Bayes, Complement Naive Bayes, Decision Tree, and Gradient Boosting might be chosen in the context of predicting child malnutrition depending on how well they performed on the data.
- **Model Calibration:**
It's important to adjust models in pediatric hunger situations to get accurate chance estimates, which can be very important for making decisions like putting kids in health prevention programs.

CHAPTER 4

ANALYSIS RESULT & DISCUSSION

Among the eight methods, we used logistic regression. In addition to logistic regression, seven additional machine learning methods have been used one after the other. We implemented various algorithms, including Random Forest, Support Vector Machine, KNearest Neighbors, Bernolli Naive Bayes, Complement Naive Bayes, Decision Tree, and Gradient Boosting. The LR strategy achieved the highest accuracy of 95.349% with a test size of 0.3%, followed by the KNN approach with an accuracy of 93.023% with a test size of 0.3 .

No	ML Model	Train Accuracy	Test Accuracy	Accuracy Score
0	LR	1.000	0.953	95.349
1	KNN	0.985	0.930	93.023
2	GBC	1.000	0.919	91.860
3	RF	1.000	0.907	90.698
4	SVM	0.955	0.895	89.535
5	CLF	1.000	0.884	88.372
6	BNB	0.804	0.791	79.070
7	CNB	0.794	0.779	77.907

Table 4.1: Results all model step by step but accuracy score is ascending

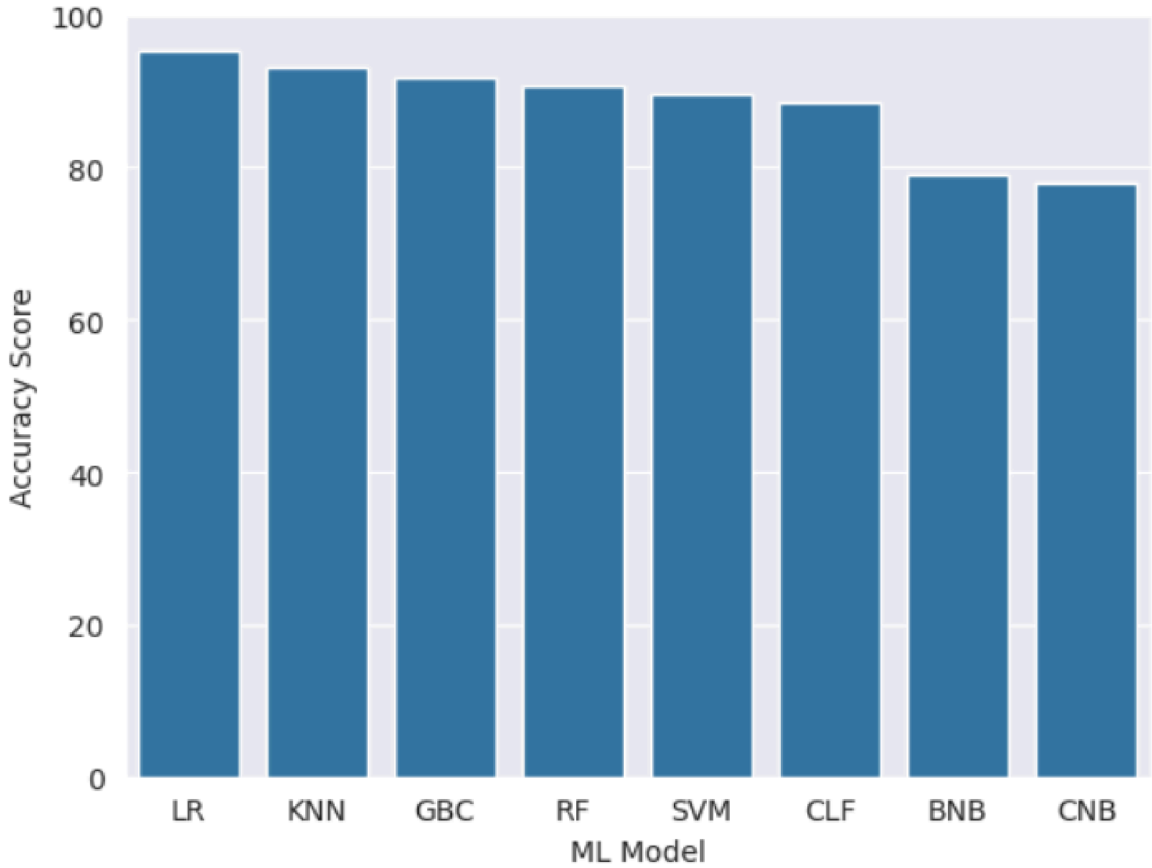

Figure 4.1: The accuracy score of our ML algorithms

The KNearest Neighbors algorithm yields identical results to the Logistic Regression classifier (Figure 4.1), with a train accuracy score of 100% and 98.5% , albeit with a different accuracy score. The KNearest Neighbors achieves 98.5% accuracy on training data and 93.02% accuracy on testing data, while LR performs exceptionally well on both training and testing data. The training and testing accuracy of LR are 100% and 95.34%, respectively. Therefore, we can infer that Logistic Regression outperforms KNearest Neighbors in terms of train and test accuracy (Figure 4.4).

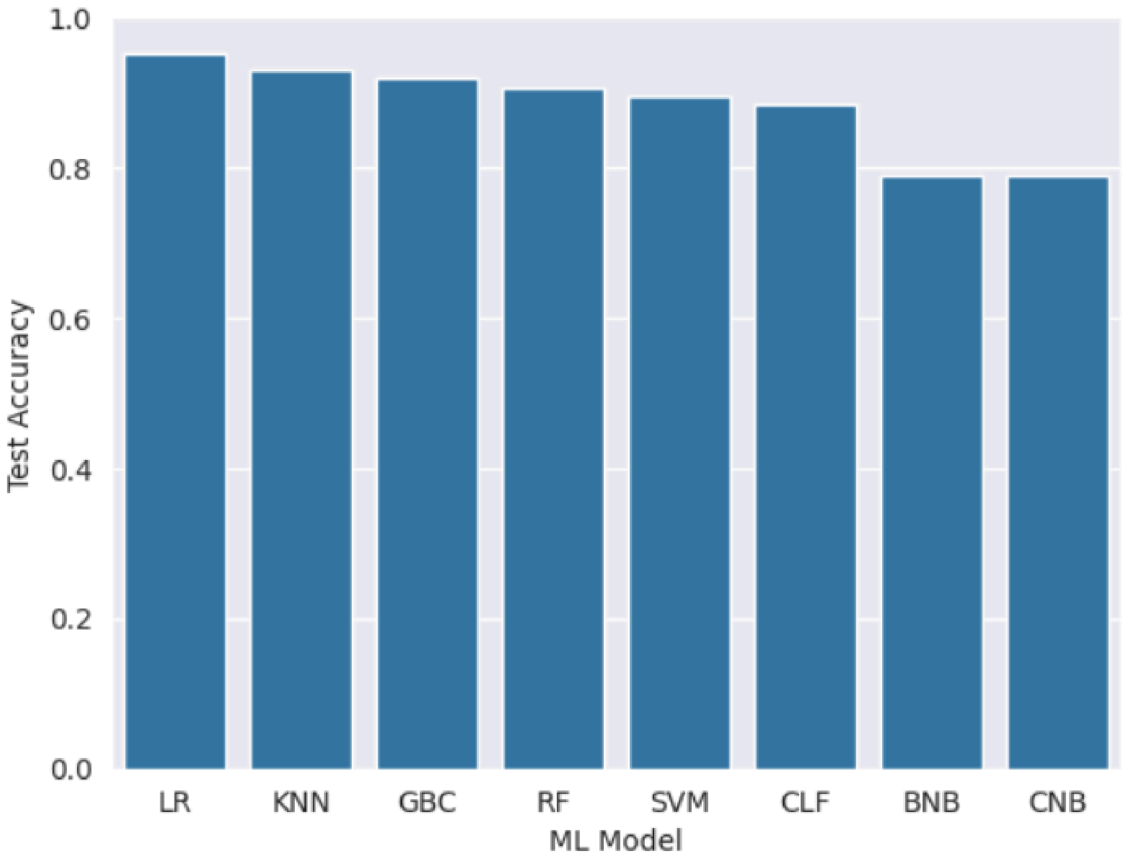

Figure 4.2: The test accuracy score of our ML algorithms

The above graph (Figure 4.1, Figure 4.2) shows us that the Gradient Boosting (GBC),Random Forest(RF),Support Vector Machine (SVM) and Decision Tree (CLF) classifiers perform moderately well with a 91.86%,90.69%,89.53% and 88.37% accuracy score. The accuracy of the training and testing data was highly comparable. Bernoulli Naive Bayes (BNB) accuracy scores are 79.07%. Complement Naive Bayes (CNB) gives a poor train or test accuracy score of 79.4% or 77.90%. Therefore, the accuracy score of CNB is 77.90%.

4.1. Classification Report and Confusion Matrix of Our Model:

- **Classification Report and Confusion Matrix for (RF):**

	Precision	Recall	F1-score	Support
Overweight	0.57	0.67	0.62	6
Sunting	0.93	0.96	0.94	69
Underweight	1.00	0.80	0.89	10
Wasting	0.00	0.00	0.00	1
Accuracy			0.91	86
Macro avg	0.63	0.61	0.61	86
Weighted avg	0.90	0.91	0.90	86

Table 4.2 : Classification Report

4	2	0	0
3	66	0	0
0	2	8	0
0	1	0	0

Table 4.3 : Confusion Matrix

- **Classification Report and Confusion Matrix for (SVM):**

	Precision	Recall	F1-score	Support
Overweight	0.57	0.67	0.62	6
Sunting	0.92	0.96	0.94	69
Underweight	1.00	0.70	0.82	10
Wasting	0.00	0.00	0.00	1
Accuracy			0.90	86
Macro avg	0.62	0.58	0.59	86
Weighted avg	0.89	0.90	0.89	86

Table 4.4 : Classification Report

4	2	0	0
3	66	0	0
0	3	7	0
0	1	0	0

Table 4.5 : Confusion Matrix

- **Classification Report and Confusion Matrix for (K-NN):**

	Precision	Recall	F1-score	Support
Overweight	0.67	0.67	0.67	6
Sunting	0.94	0.97	0.96	69
Underweight	1.00	0.90	0.95	10
Wasting	0.00	0.00	0.00	1
Accuracy			0.93	86
Macro avg	0.65	0.63	0.64	86
Weighted avg	0.92	0.93	0.92	86

Table 4.6 : Classification Report

4	2	0	0
2	67	0	0
0	1	9	0
0	1	0	0

Table 4.7 : Confusion Matrix

- **Classification Report and Confusion Matrix for (LR):**

	Precision	Recall	F1-score	Support
Overweight	0.80	0.67	0.73	6
Sunting	0.96	0.99	0.97	69
Underweight	1.00	1.00	1.00	10
Wasting	0.00	0.00	0.00	1
Accuracy			0.95	86
Macro avg	0.69	0.66	0.67	86
Weighted avg	0.94	0.95	0.95	86

Table 4.8 : Classification Report

4	2	0	0
1	68	0	0
0	0	10	0
0	1	0	0

Table 4.9 : Confusion Matrix

- **Classification Report and Confusion Matrix for (BNB):**

	Precision	Recall	F1-score	Support
Overweight	0.00	0.00	0.00	6
Sunting	0.80	0.99	0.88	69
Underweight	0.00	0.00	0.00	10
Wasting	0.00	0.00	0.00	1
Accuracy			0.79	86
Macro avg	0.20	0.25	0.22	86
Weighted avg	0.64	0.79	0.71	86

Table 4.10 : Classification Report

0	6	0	0
1	68	0	0
0	10	0	0
0	1	0	0

Table 4.11 : Confusion Matrix

- **Classification Report and Confusion Matrix for (CNB):**

	Precision	Recall	F1-score	Support
Overweight	0.29	1.00	0.44	6
Sunting	0.93	0.80	0.86	69
Underweight	1.00	0.60	0.75	10
Wasting	0.00	0.00	0.00	1
Accuracy			0.78	86
Macro avg	0.55	0.60	0.51	86
Weighted avg	0.88	0.78	0.81	86

Table 4.12 : Classification Report

6	6	0	0
14	55	0	0
0	4	6	0
1	0	0	0

Table 4.13 : Confusion Matrix

- **Classification Report and Confusion Matrix for (CLF):**

	Precision	Recall	F1-score	Support
Overweight	0.33	0.33	0.33	6

	Precision	Recall	F1-score	Support
Sunting	0.92	0.94	0.93	69
Underweight	1.00	0.90	0.95	10
Wasting	0.00	0.00	0.00	1
Accuracy			0.88	86
Macro avg	0.56	0.54	0.55	86
Weighted avg	0.87	0.88	0.88	86

Table 4.14 : Classification Report

2	4	0	0
4	65	0	0
0	1	9	0
0	1	0	0

Table 4.15 : Confusion Matrix

- **Classification Report and Confusion Matrix for (GB):**

	Precision	Recall	F1-score	Support
Overweight	0.75	0.50	0.60	6
Sunting	0.92	0.99	0.95	69
Underweight	1.00	0.80	0.89	10
Wasting	0.00	0.00	0.00	1
Accuracy			0.92	86
Macro avg	0.67	0.57	0.61	86
Weighted avg	0.91	0.92	0.91	86

Table 4.16 : Classification Report

3	3	0	0
1	68	0	0
0	2	8	0
0	1	0	0

Table 4.17 : Confusion Matrix

4.2. RESULT DISCUSSION:

The child malnutrition prediction model's outcomes show how well it can identify children who are at risk based on characteristics including age, height, weight, access to healthcare, and food availability. Precision, recall, and the F1 score are important performance indicators that indicate the model successfully strikes a compromise between detecting malnourished children and reducing false positives. The precision-recall trade-off, however, emphasizes the need to give careful thought to situations where failing to detect a hungry kid (false negative) is more serious than misidentifying a healthy youngster (false positive). According to feature significance analysis, characteristics such as food availability and weight-for-age are powerful predictors that may help direct public health initiatives to address nutritional shortfalls and healthcare access in disadvantaged communities. Issues such as data imbalance and context-specific changes limit the model's broad applicability, underscoring the need for more comprehensive and targeted data collection. Finally, the model offers useful insights for policy formation and resource allocation in the battle against childhood malnutrition.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT & SUSTAINABILITY:

5.1. IMPACT ON SOCIETY:

The impact of child malnutrition on society is significant and wide-ranging. Malnourished children have an increased risk of death, developmental delays, and chronic health problems. This not only has an impact on individuals' quality of life, but it also perpetuates a cycle of poverty and underdevelopment among communities. Malnourished children often have cognitive deficits, which contribute to low academic performance and restricted future career chances. Malnutrition has several negative societal effects, such as higher medical expenses, lower economic output, and more demand for social services. Addressing child malnutrition is critical for breaking the cycle, encouraging healthy growth, and ensuring that future generations can constructively contribute to society.

5.2. IMPACT ON ENVIRONMENT:

The environment significantly influences the frequency of child malnutrition, and hunger exacerbates environmental deterioration. Poor farming methods, deforestation, and unsustainable land use all contribute to food insecurity, which has a direct influence on child nutrition. Environmental issues such as soil erosion, water shortages, and climate change further diminish crop yields and the availability of a variety of nutrient-dense foods. Societies that heavily rely on natural resources for survival may experience an escalation in malnutrition. We need long-term, better control of natural resources and permanent farming methods to increase food security, reduce the number of malnourished children, and promote a healthy climate.

5.2. SUSTAINABILITY PLAN:

In order to combat child malnutrition, a comprehensive sustainability strategy must include policy interventions, education, agriculture, and health. The essential elements consist of:

1. **Enhancing Food Security:** Sustainable farming, crop variety, and equal access to healthy food sources are key to fighting malnutrition. Supporting small-scale farmers, minimizing food waste, and improving food delivery networks are all critical measures in the sustainability strategy.

2. **Improving Healthcare Access:** Providing frequent health screenings, immunizations, and nutritional supplements to disadvantaged populations aids in the early diagnosis and intervention of malnourished children. Mobile clinics and community health workers may increase healthcare access, especially in rural and underserved regions.

3. **Education and awareness:** Educating communities about nutrition, cleanliness, and the benefits of breastfeeding may help lower malnutrition rates. Improving maternal nutrition and educating parents and caregivers about nutrition should be the main goals of programs.

4. **Policy and Governance:** Governments must establish and execute policies to reduce poverty, increase food security, and improve healthcare infrastructure. When it comes to making sure initiatives last, the public and private sectors must work together.

5. **Climate Resilience:** Implementing climate-resilient agricultural methods to deal with environmental changes may help ensure food production in the face of climate unpredictability. Training farmers to use environmentally friendly and resource-efficient agricultural practices will help ensure long-term sustainability.

This multi-sectoral strategy guarantees that initiatives to address child malnutrition are long-lasting, socially rewarding, and ecologically responsible.

CHAPTER 6

CONCLUSION & FUTURE STUDY:

6.1. CONCLUSION:

In this research, we examined the eight machine learning algorithms in terms of accuracy scores, training accuracy, and test accuracy. This was done to enable the algorithms to accurately determine whether a child is malnourished. Among these algorithms, logistic regression and random forest produced the most accurate predictions in Asia's dataset. For this study, we effectively deployed our machine learning algorithms throughout Asia. We use these algorithms to determine whether a child is undernourished. Our first objective is to maintain control of the situation. They will therefore initiate various initiatives aimed at educating people about this issue and assisting them in becoming more reasonable. They will provide the necessary guidance to address this issue effectively. Furthermore, in the future, it will be crucial to distinguish between children who are malnourished and those who are not. This will eventually lead to the provision of wholesome food, necessary medications, and other beneficial health supplements. As a result, our approach will be useful in managing the situation and eliminating child malnutrition.

6.2. FUTURE STUDY:

In Bangladesh, ongoing research on child malnutrition should prioritize the resolution of the distinctive socio-economic, environmental, and healthcare obstacles that affect vulnerable populations. Our next research will focus on Bangladeshi hilly districts and coastal areas, which have poor education rates and a significant risk of child marriage. This is because in this area, children under the age of five are at the highest risk of malnutrition. Comprehensive data collection in rural and underprivileged regions is critical for enhancing prediction model accuracy, especially in terms of maternal health, food security, and healthcare access. Researchers should prioritize region-specific studies that examine local eating habits, cultural norms, and economic situations to implement changes that are most effective in those areas. Furthermore, given Bangladesh's sensitivity to climate change, it is crucial to include environmental variables such as floods, agricultural yields, and water availability in order to understand how climate-induced food poverty impacts child nutrition. Ethical issues, such as ensuring that prediction models are free of biases that may prejudice marginalized groups, must also be considered. In order to maximize intervention techniques and guarantee long-lasting gains in child health and nutrition outcomes, it would be helpful to assess the long-term efficacy of Bangladesh's nutrition programs and policies, particularly those that are based on data-driven approaches.

REFERENCE

[1]. Kar, S., Pratihar, S., Nayak, S., Bal, S., Gururai, H. L., & Ravikumar, V. (2021, December). Prediction of child malnutrition using machine learning. In *2021 10th international conference on internet of everything, microwave engineering, communication and networks (IEMECON)* (pp. 01-04). IEEE.

[2]. Talukder, A., & Ahammed, B. (2020). Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. Nutrition, 78, 110861.

[3]. Ula, M., Ulva, A. F., Saputra, I., Mauliza, M., & Maulana, I. (2022). Implementation of machine learning using the k-nearest neighbor classification model in diagnosing malnutrition in children. *Multica Science and Technology*, *2*(1), 8-13.

[4]. Bitew, F. H., Sparks, C. S., & Nyarko, S. H. (2022). Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia. Public health nutrition, 25(2), 269-280.

[5]. Rahman, S. J., Ahmed, N. F., Abedin, M. M., Ahammed, B., Ali, M., Rahman, M. J., & Maniruzzaman, M. (2021). Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach. Plos one, 16(6), e0253172.

[6]. Rachmi, C. N., Agho, K. E., Li, M., & Baur, L. A. (2016). Stunting, underweight and overweight in children aged 2.0–4.9 years in Indonesia: prevalence trends and associated risk factors. *PloS one*, *11*(5), e0154756.

[7]. Qasrawi, R., Sgahir, S., Nemer, M., Halaikah, M., Badrasawi, M., Amro, M., ... & Al-Jawaldeh, A. (2024). Machine Learning Approach for Predicting the Impact of Food Insecurity on Nutrient Consumption and Malnutrition in Children Aged 6 Months to 5 Years. *Children*, *11*(7), 810.

[8]. Chilyabanyama, O. N., Chilengi, R., Simuyandi, M., Chisenga, C. C., Chirwa, M., Hamusonde, K., ... & Bosomprah, S. (2022). Performance of machine learning classifiers in classifying stunting among under-five children in Zambia. *Children*, *9*(7), 1082.

[9]. Hemo, S. A., & Rayhan, M. I. (2021). Classification tree and random forest model to predict under-five malnutrition in Bangladesh. *Biom Biostat Int J*, *10*(3), 116-123.

[10]. Thangamani, D., & Sudha, P. (2014). Identification of malnutrition with use of supervised datamining techniques–decision trees and artificial neural networks. *Int J Eng Comput Sci*, *3*(09).

[11]. Kawo, M. A., Idris, M. N., Haruna, K., & Kaita, F. U. (2024). PREDICTING UNDERNUTRITION RISK FACTORS USING MACHINE LEARNING TECHNIQUES IN NIGERIAN UNDER FIVE CHILDREN.

[12]. Aryuni, M., Miranda, E., Kumbangsila, M., Zakiyyah, A. Y., Sano, A. V. D., & Bhatti, F. M. (2023, April). Comparison of Nutritional Status Prediction Models of Children Under 5 Years of Age Using Supervised Machine Learning. In *Proceeding of the 3rd International Conference on Electronics, Biomedical Engineering, and Health Informatics: ICEBEHI 2022, 5–6 October, Surabaya, Indonesia* (pp. 265-277). Singapore: Springer Nature Singapore.

[13]. Shinde, M. A. A., & Sahasrabuddhe, D. V. Diagnosing Malnutrition in Preschool Children Using Machine Learning Approach: A Review.

[14]. Turjo, E. A., & Rahman, M. H. (2024). Assessing risk factors for malnutrition among women in Bangladesh and forecasting malnutrition using machine learning approaches. *BMC nutrition*, *10*(1), 22.

[15]. Nirmani, H. M. C., & Kudagamage, U. P. (2024, February). Ensemble Approach for Early Prediction of Malnutrition Level of Children: A Case Study on Children Under Five Years Old. In *2024 4th International Conference on Advanced Research in Computing (ICARC)* (pp. 91-96). IEEE.

[16]. Shahriar, M. M., Iqubal, M. S., Mitra, S., & Das, A. K. (2019, July). A Deep Learning Approach to Predict Malnutrition Status of 0-59 Month's Older Children in Bangladesh. In *2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)* (pp. 145-149). IEEE.

[17]. S. Das and J. Gulshan, "Different forms of malnutrition among under five children in Bangladesh: a cross sectional study on prevalence and determinants", BMC Nutrition, vol. 3, no. 1, 2017.

[18]. K. Bhowmik and S. Das, "On exploring and ranking risk factors of child malnutrition in Bangladesh using multiple classification analysis", BMC Nutrition, vol. 3, no. 1, 2017.

[19]. "Nutrition country profiles: Bangladesh summary", Fao.org, 2010. [Online]. Available: http://www.fao.org/ag/agn/nutrition/bgd_en.stm.

[20]. Black RE, Allen LH, Bhutta ZA, Caulfield LE, De Onis M, Ezzati M, et al. Maternaland child undernutri tion: global and regional exposures and health consequences. Lancet. 2008; 371(9608):243–60.

[21]. Black RE, Victora CG, WalkerSP,BhuttaZA,Christian P, De OnisM, et al.Maternal and childundernu trition and overweight in low-income and middle-income countries. Lancet. 2013; 382(9890):427–51.

[22]. Mitra, M. (2015). Stunting problems and interventions to prevent stunting (a Literature Review). Jurnal Kesehatan Komunitas, 2(6), 254-261.

[23]. Widjayatri, R. D., Fitriani, Y., & Tristyanto, B. (2020). Sosialisasi Pengaruh Stunting Terhadap Pertumbuhan dan Perkembangan Anak Usia Dini. Murhum: Jurnal Pendidikan Anak Usia Dini, 16-27.

[24]. Waliulu, S. H., Ibrahim, D., & Umasugi, M. T. (2018). Pengaruh edukasi terhadap tingkat pengetahuan dan upaya pencegahan stunting anak usia balita. Jurnal Penelitian Kesehatan" SUARA FORIKES"(Journal of Health Research" Forikes Voice"), 9(4), 269- 272.

[25]. Supariasa, I. D. (2001). Penilaian status gizi. Jakarta: Penerbit Buku. Kedokteran EGC.

[26]. Ula, M., Hendriana, Y., & Hardi, R. (2016, October). An expert system for early diagnose of vitamins and minerals deficiency on the body. In 2016 international conference on information technology systems and innovation (ICITSI) (pp. 1-6). IEEE.

[27]. Primasari, Y., & Keliat, B. A. (2020). Parenting Practices as Prevention of Stunting Impact in Children's Psychosocial Development. Jurnal Ilmu Keperawatan Jiwa, 3(3), 263-272.

[28]. Ngiam KY, Khor W. Big data and machine learning algorithms for health care delivery. Lancet Oncol. 2019;20(5):262–73.

[29]. Dutta, M., Hasan, K. M. M., Akter, A., Rahman, M. H., & Assaduzzaman, M. (2024). An interpretable machine learning-based breast cancer classification using XGBoost, SHAP, and LIME. *Bulletin of Electrical Engineering and Informatics*, *13*(6), 4306-4315.

[30]. Rabby, M. H., Islam, O., Assaduzzaman, M., & Dutta, M. (2023, December). Tuberculosis Disease Detection from Chest X-rays Using Deep Learning Techniques. In *2023 26th International Conference on Computer and Information Technology (ICCIT)* (pp. 1-6). IEEE

[31]. Dutta, M., Islam Sujan, M. R., Mojumdar, M. U., Chakraborty, N. R., Marouf, A. A., Rokne, J. G., & Alhajj, R. (2024). Rice Leaf Disease Classification—A Comparative Approach Using Convolutional Neural

Network (CNN), Cascading Autoencoder with Attention Residual U-Net (CAAR-U-Net), and MobileNet-V2 Architectures. *Technologies*, *12*(11), 214.

[32]. Assaduzzaman, M., Dutta, M., Saha, A., & Paul, S. G. (2024). ALSA-3: Customized CNN model through ablation study for Alzheimer's disease classification. *Informatics in Medicine Unlocked*, *50*, 101584.

Child Malnutrition.docx-3.pdf

ORIGINALITY REPORT

25%	19%	16%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	5%
2	Submitted to Daffodil International University Student Paper	2%
3	journal.universitasmulia.ac.id Internet Source	1%
4	publikasi.lldikti10.id Internet Source	1%
5	ijeijournal.com Internet Source	1%
6	Submitted to University of Newcastle upon Tyne	1%