"Prediction of Child Malnutrition Using AI/ML Algorithm"

[1]Manaswini P, [2]Adarsh S, [3]Parikshith Samanth, [4]Sai Srusti A S
Information Science and Engineering, Malnad College of Engineering
Hassan-573202, India
Email id: manaswiniprakash0@gmail.com, abj@mcehassan.ac.in, adarsh1308.s@gmail.com, parikshithsamanth@gmail.com, gowdasrusti5@gmail.com

***Abstract*— Synthetic Intelligence (AI) and machine learning (ML) are hastily transforming numerous fields, imparting advanced and faster outcomes. As facts volume grows, leveraging AI and ML becomes crucial for reading big datasets correctly. This paper surveys the software of ML techniques in predicting baby malnutrition, a crucial thing in allowing early interventions and prioritizing remedies. Key clinical factors which includes age, gender, height, and weight are utilized in building predictive fashions. Current research employs supervised getting to know algorithms like SVM, Linear Regression, selection bushes, AdaBoost, and Random wooded area, reaching accuracies of 85% to 95%. But, maximum models lack real-time applicability.**

Keywords— *Data Science, Malnutrition, Child, Machine Learning, AI, Research, Existing Works, Algorithms*

I. INTRODUCTION

Right nutrients is important for the survival, growth, and improvement of youngsters, yet malnutrition remains a vast global fitness difficulty. Key factors influencing malnutrition encompass age, gender, top, weight, Weight-for-Age Z-rating (WAZ), peak-for-Age Z-score (HAZ), and Weight-for-top Z-rating (WHZ)[4][7]. Modern manual techniques for malnutrition evaluation are often misguided, main to behind schedule interventions and adverse fitness results. Early detection is vital for presenting centered treatments, enhancing child health, and safeguarding a country's destiny personnel, which immediately affects economic boom. Device learning (ML) has emerged as a powerful tool within the scientific field, supplying advanced accuracy and performance in prediction obligations. With the growing availability of huge datasets, ML techniques provide valuable insights and aid selection-making strategies. A number of the 3 primary types of ML—supervised, unsupervised, and semi-supervised—supervised learning is specifically powerful for malnutrition prediction[10]. This paper explores present studies, discusses various algorithms which include help Vector Machines (SVM), choice bushes, Random wooded area, and AdaBoost, and highlights their potential for early and accurate malnutrition detection.

II. RELATED WORKS

In this section, we summarize and examine current research related to infant malnutrition prediction the usage of AI/ML algorithms. These researches examine the numerous algorithms, datasets, and methodologies used to expect and classify child malnutrition.

A. *Paper Title: Efficient Machine Learning for Malnutrition Prediction among under-five children in India Authors: Saksham Jain, Tayyibah Khanam, Ali Jafar Abedi,Abid Ali Khan Year: 2022*

Description: Infant malnutrition is taken into consideration to be one of the main causes of infant mortality and malnutrition. This take a look at turned into aimed to leverage the advantages supplied by way of device gaining knowledge of models in phrases of figuring out and accurately predicting great elements of malnutrition. For this look at, the kids's recode files from the Indian Demographic and fitness Survey (IDHS) datasets from 2005-2006 and 2015-2016 were used. To have a look at the nutritional status of youngsters aged 0-59 months, this have a look at seems at stunting (top-for-age), wasting (Weight-for height), and concurrent stunted losing (peak-for-age-Weight for-top). Ordinary device mastering models, Tabular Deep getting to know frameworks, H2O base fashions, and automobile ML models are the 4 sorts of machine studying fashions hired in our research. This research located that automated system getting to know algorithms and Tabular Deep gaining knowledge of frameworks.

Methodology: Regular Machine Learning models, Tabular Deep Learning frameworks, H2O base models, and AutoML models

Limitations: Existing studies on child malnutrition prediction the usage of AI/ML algorithms have boundaries, including using smaller datasets that may not absolutely constitute the trouble. Those models are also often now not suitable for actual-time packages and keep in mind simplest a restricted set of factors, which may also affect the accuracy of predictions.

Citation: [1]. "Efficient Machine Learning for Malnutrition Prediction among under-five children in India",\Saksham Jain, Tayyibah Khanam, Ali Jafar Abedi,Abid Ali Khan, 2022.

B. *Paper Title: Ensemble Approach for Early Prediction of Malnutrition Level of Children: A Case study on Children under Five Years Old Authors: H.M.C. Nirmani, U.P. Kudagamage Year: 2024*

Description: Malnutrition, an international health difficulty affecting youngsters, leads to extreme brief- and lengthy-time period results on their boom and improvement, in particular within the essential 1–60 months age range. This observe aims to increase a predictive version for early detection of malnutrition the usage of ensemble mastering strategies, based on data from 574 kids underneath five inside the

Lunugala location. After information preprocessing, the dataset was cut up into 70% education and 30% checking out sets, and nine gadget gaining knowledge of algorithms, consisting of SVM, selection Tree, Random woodland, and XGBoost, had been evaluated.

Methodology: Support Vector Machine (SVM), Logistic Regression (LR), AdaBoost and Extreme Gradient Boosting (XGBoost) were used.

Limitations: Many studies focus entirely on constructing ML models for toddler malnutrition prediction, however these fashions are regularly not well matched with actual-time use. In addition they tend to rely upon smaller datasets, which can limit the robustness of the predictions. Moreover, in some instances, the evaluation of the ML fashions is not performed very well, main to an incomplete know-how in their overall performance and capability obstacles.

Citation: [2]. "Ensemble Approach for Early Prediction of Malnutrition Level of Children: A Case study on Children under Five Years Old", H.M.C. Nirmani, U.P. Kudagamage, 2024.

C. *Paper title: Implementation of Hybrid Bat Algorithm-Ensemble on Side Effect Prediction: Case Study Metabolism and Nutrition Disorders Authors: Dzaky Raihan Ahmad, Jondri, Isman Kurniawan Year: 2024*

Description: Drug aspect effects, ranging from moderate issues like nausea to intense outcomes, pose vast health dangers. Contemporary detection methods, reliant on scientific trials, are luxurious and often omit capability results. This examine addresses those challenges through the usage of a Hybrid Bat algorithm (HBA) for feature choice and an ensemble technique combining Random forest, AdaBoost, and XGBoost to are expecting drug side effects, that specialize in metabolism and nutrients issues.

Methodology: Hybrid Bat Algorithm (HBA), Random Forest, AdaBoost, and XGBoost

Limitations: A few studies goal to expect drug side effects, with a focus on metabolism and vitamins issues, but they're unable to as it should be expect child malnutrition. These fashions regularly generate much less correct consequences, as they're now not specifically designed for malnutrition prediction, proscribing their applicability and effectiveness in addressing infant malnutrition issues.
Citation: [3]. "Implementation of Hybrid Bat Algorithm-Ensemble on Side Effect Prediction: Case Study Metabolism and Nutrition Disorders", Dzaky Raihan Ahmad, Jondri, Isman Kurniawan, 2024.

D. *Paper Title: Machine Learning Approaches for Prediction of Nutrition Deficiency among Women of Different Age Groups Authors: Javeria Ali , Waseemullah, Masood Ahmed Khan ,Najeed Ahmed Khan Year: 2022*

Description: Malnutrition is a sizable fitness problem among girls in underdeveloped areas, in particular in Pakistan, due to constrained healthcare access. This take a look at aims to predict nutritional deficiencies the usage of system getting to know models like SVM, Random forest, Logistic Regression, KNN, and Naïve Bayes. Outcomes display Random woodland accomplished the best accuracy, making it the maximum dependable for future predictions. Testing on similar datasets showed the model's reliability for assessing women's health conditions.

Methodology: Support vector machine, random forest, Logistic Regression, KNN algorithm used

Limitations: Some studies focus on predicting nutrition deficiencies for women, with age-based constraints, making them unsuitable for child malnutrition prediction. Additionally, the datasets used in these studies are often incompatible with the needs of child malnutrition models, limiting their relevance and accuracy for this specific issue.
Citation: [4]. "Machine Leaning Approaches for Prediction of nutrition deficiency among Women of different Age Groups", Javeria Ali , Waseemullah, Masood Ahmed Khan ,Najeed Ahmed Khan , 2022.

III. METHODOLOGY

Data Science: Data Science is a process of reading data from extraordinary prospective s and extraction of useful facts from the processed facts. Information science applied on n quantity of fields and used to resolve real international troubles. Records technology helps many strategies.

Classification Rules (Classifiers): Basically class is used to categorise each object in a fixed of information into one of the predefined set of training or businesses. Class methods make use of mathematical strategies for problem fixing

Ex: worker statuses in a employer (leaves or stay)

To predict which cutting-edge employees are possibly to leave inside the destiny.

Inside the project we use both "naive bayes" and "KNN" or "choice tree" classifier to technique preceding records and for prediction. Those special algorithms are most efficient and takes less time for processing records. These algorithms work great for n wide variety of parameters. These algorithms help all formats of facts.

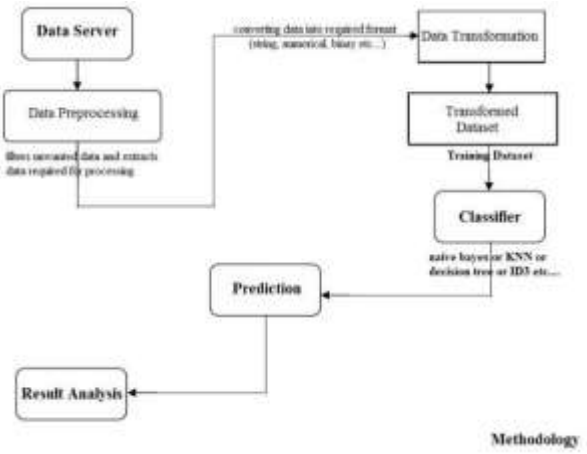


Fig. 3.1. Methodology

IV. ALGORITHMS

4.1.1 SVM Algorithm:

Support Vector Algorithm (SVM) is a effective device gaining knowledge of set of rules broadly used for both linear

and nonlinear category, as well as regression and outlier detection tasks[5]. SVMs are particularly adaptable, making them suitable for numerous packages together with textual content class, photograph type, junk mail detection, handwriting identification, gene expression evaluation, face detection, and anomaly detection[6].

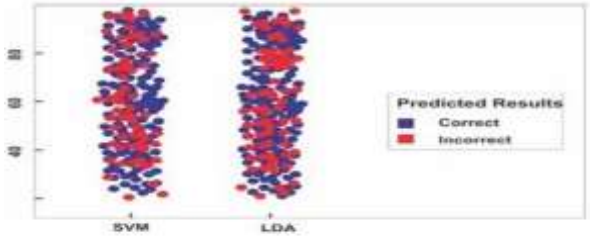
The equation for the linear hyperplane can be written as:

$$w^T x + b = 0$$

The vector W represents the normal vector to the hyperplane. i.e the direction perpendicular to the hyperplane. The parameter b in the equation represents the offset or distance of the hyperplane from the origin along the normal vector w. The distance between a data point x_i and the decision boundary can be calculated as:

$$d_i = (w^T x_i + b) / \|w\|$$

where ||w|| represents the Euclidean norm of the weight vector w.



Predicted results using SVM and LDA

Fig. 4.1.1. Predicted results

4.1.2 Regression Algorithm

The Linear Regression algorithm involves the implementation of a system primarily based at the least squares method[7]. This pseudocode outlines the Linear Regression technique, facilitating its implementation in high-stage programming languages.

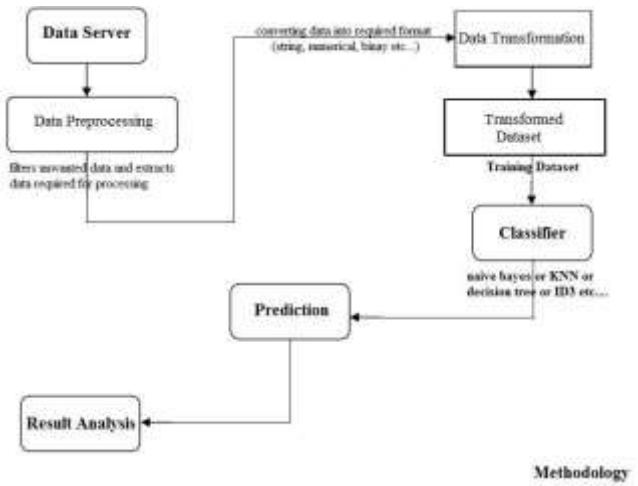


Fig. 4.1.2. Methodology of Linear Regression

Mathematical formula:
Simple linear regression:

This is the simplest form of linear regression, and it involves only one independent variable and one dependent variable. The equation for simple linear regression is:

$$y = \beta_0 + \beta_1 X$$

where, Y is the dependent variable

X is the independent variable

β_0 is the intercept and β_1 is the slope

Multiple linear regression:

This involves more than one independent variable and one dependent variable.

The equation for multiple linear regression is:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n$$

where, Y is the dependent variable

$X_1, X_2, ..., X_n$ are the independent variables

β_0 is the intercept and $\beta_1, \beta_2, ..., \beta_n$ are the slopes

4.1.3 Decision Tree Algorithm

Decision trees are flexible and effective flowchart-like models suitable for both prediction and classification obligations. In contrast to linear regression models, which anticipate a linear courting among predictors and objectives, Decision tree perform without such constraints[10]. As non-parametric fashions, they could correctly capture complicated nonlinear relationships between predictors and goals.

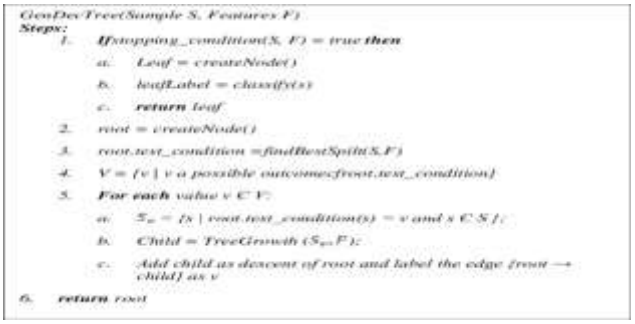


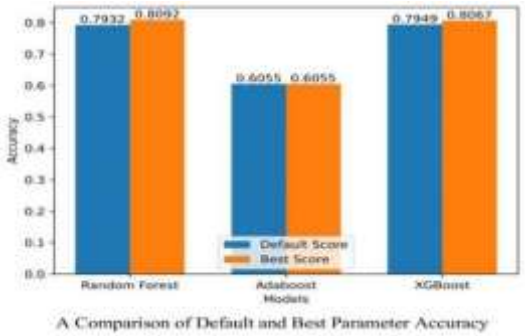
Fig. 4.1.3. Algorithm for decision tree

Mathematical formula:
Information Gain = Entropy(S)- [(Weighted Avg) *Entropy(each feature)]
Entropy: Entropy is a metric to measure the impurity in a given attribute. It specifies randomness in data. Entropy can be calculated as:

Entropy(s)= -P(yes)log2 P(yes)- P(no) log2 P(no)

where, S= Total number of samples
P(yes)= probability of yes
P(no)= probability of no

V. RESULTS



A Comparison of Default and Best Parameter Accuracy

Fig. 5.1. A comparison of Default and Best Parameter accuracy

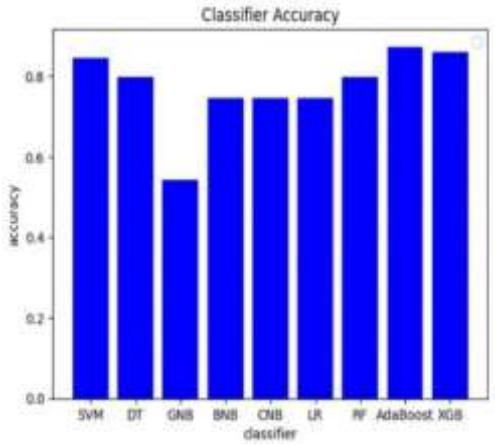


Fig. 5.2. Classifier accuracy

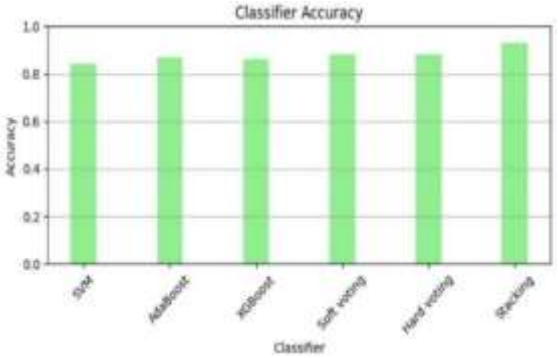


Fig. 5.3. Classifier accuracy

	TPR (Sensitivity or Recall)	FPR (Fall out)	Accuracy
Ideal Values	*100%*	*0 %*	*100 %*
SVM (Proposed)	0.69	0.5	0.77
LDA (Proposed)	0.54	0.46	0.59
MLR Model 1 [31]	0.74	0.717	0.73
SVM Model 1 [31]	0.63	0.740	0.68
MLR Model 2 [31]	0.72	0.761	0.74
SVM Model 2 [31]	0.63	0.739	0.68

Fig. 5.4. A table showing comparison of models based on TPR, FPR and accuracy

VI. CONCLUSION

Predicting malnutrition in kids at early stages performs a vital position in proper remedies and child growth. Prediction of baby malnutrition relies upon on many clinical elements consisting of age of the kid, gender, top, weight and many others. There are many research-works in this child malnutrition prediction. Few topics simply provided concept and few works accomplished the implementation the usage of ML models and outcomes proven. All existing models simply suits for static records and suitable for experiments but no longer suitable for actual time.

VII. REFERENCES

[1]. “Efficient Machine Learning for Malnutrition Prediction among under-five children in India”, Saksham Jain, Tayyibah Khanam, Ali Jafar Abedi,Abid Ali Khan, 2022
[2]."Ensemble Approach for Early Prediction of Malnutrition Level of Children: A Case study on Children under Five Years Old", H.M.C. Nirmani, U.P. Kudagamage, 2024
[3]. "Implementation of Hybrid Bat Algorithm-Ensemble on Side Effect Prediction: Case Study Metabolism and Nutrition Disorders”, Dzaky Raihan Ahmad, Jondri, Isman Kurniawan, 2024
[4]. “Machine learning approaches for Prediction of Nutrition Deficiency among Women of Different Age Groups", Javeria Ali, Waseemullah, Masood Ahmed Khan ,Najeed Ahmed Khan , 2022
[5]. “Prediction and detection of nutrition deficiency using machine learning", Amit Kumar Mishra, Neha Tripathi, Ashish Gupta, Deepak Upadhyay, Neeraj Kumar Pandey, 2023
[6]. “Malnutrition: Afghanistan’s silent emergency”, Retrieved June 13, 2019
[7]. Ministry of Public Health, Afghanistan. (2019). "Afghanistan Nutrition Cluster, 2018 annual report”.
[8]. Ministry of Public Health, Afghanistan. (2018). “Integrated Management of Acute Malnutrition National Guidelines”. Kabul, Afghanistan: Public Nutrition Directorate.
[9]. Khare, S., Kavyashree, S., Gupta, D., & Jyotishi, A. (2017). "Investigation of Nutritional Status of Children based on Machine Learning Techniques using Indian Demographic and Health Survey Data”. Procedia Computer Science, 115(2017), 338–349.
[10]. Akseer, N., Bhatti, Z., Mashal, T., Soofi, S., Moineddin, R., Black, R. E., & Bhutta, Z. A. (2018). "Geospatial inequalities and determinants of nutritional status among women and children in Afghanistan”: an observational study. The Lancet GlobalHealth, 6(4), e447–e459.