Article

Machine learning model for predicting fetal nutritional status

B. Selemani, **D. Machuve**, **N. Mduma**
Nelson Mandela African Institution of Science and Technology, Arusha Tanzania
E-mail: selemanib@nm-aist.ac.tz; dina.machuve@nm-aist.ac.tz, neema.mduma@nm-aist.ac.tz

Received 25 April 2023; *Accepted 20 July 2023*; *Published online 23 October 2023*; *Published 1 March 2024*

Abstract
Malnutrition tends to be one of the most important reasons for child mortality in Tanzania and other developing countries, in most cases during the first five years of life. This research was conducted todevelop machine learning model for predicting fetal nutritional status. Several machine learning techniques such as AdaBoost, Logistic Regression, Support Vector Machine, Random Forest, Naïve Bayes, Decision Tree, K-nearest neighbor and Stochastic Gradient Descent, were used to categorize the children in the test dataset as "malnourished" or "nourished". The accuracy, sensitivity, and specificity of these algorithms' prediction abilities were comparedusing performance measures such as accuracy, sensitivity, and specificity. Results show that malnutrition status can be predicted using Random Forest machine learning technique which was about 98% and brings positive impact to the society. The study findings indicated a need for more attention on nutrition to expected mothers and children under five to be well administered with the government and the society at large by putting relevance to the suggestion that cooperation between government organizations, academia, and industry is necessary to provide sufficient infrastructure support for the future society.

Keywords malnutrition; mobile application; machine learning; Tanzania.

Computational Ecology and Software
ISSN 2220-721X
URL: http://www.iaees.org/publications/journals/ces/online-version.asp
RSS: http://www.iaees.org/publications/journals/ces/rss.xml
E-mail: ces@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

Malnutrition remains endemic in many poor nations, particularly in Sub-Saharan Africa and parts of Asia. According to UNICEF, WHO and the World Bank, Stunting affected an estimated 22.0% or 149.2 million children under five (5yrs) globally in 2020, over a third resided in Africa (World Health Organization, 2021). In Tanzania, around 3 million children under five years of age are estimated to be stunted due to inadequate nutrition. The World Health Report (2003, 2003) argued that in Mainland, the level of stunting was considered to be very high (≥30%) in 15 regions out of 26 (Tanzania Food and Nutrition Center, 2015). Tanzania National Nutrition Survey (2018, 2019) reported that stunting affect about 10.0 % of children countrywide. Whereas the most affected regions with a prevalence of stunting exceeding 40% were mentioned to be Ruvuma (41.0%),

4 Discussion

This study, however, shows that machine learning algorithms and the use of mobile application can be used to predict malnutrition based on common risk factors and life style, which can aid in the creation of human behavior and treatments to avoid malnutrition among Tanzania's children and expected mothers.

On the age of giving birth the result shows that there is a big chance to streamline and make change to the society on nutrition behaviors as to make sure that may be the government should start giving nutritional supplements to ladies of such ages all over the country.

The results on children nutrition status indicated that we are going to have a society with a big problem in nutrition so we need to make sure those who are having severe and moderate malnutrition status are helped to overcome that situation by providing them with needed balanced diet, while those who are obese and overweight were needed to follows the advice from nutritionist.

The government has to give power to social welfare officers to collaborate with nutritionist to overcome this situation and if possible, some bylaws should be put in for the parents who will refuse to follow nutrition instruction to their children.

There is need to change behavior of nutrition situation in our society. Nutritionist and social welfare officers have the mandate to spread and streamline behavioral aspect to overcome some nutrition problems which are caused by some culture and behavior.

The main goal of this study is to predict malnutrition in children under the age of five. Several well-known machine learning (ML) methods were used to achieve our goal: AdaBoost, k-nearest neighbors (k-NN), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Logistic Regression (LR), and Random Forest (RF), Decision Tree. These machine learning methods were tested on a sample of 80% of the dataset, (training dataset, n=11,649) and approved on the remaining 20% (test dataset, n=2,913). 10-fold cross validation was used to train all of the models. On the training set, we used 10-fold cross validation, and on the testing set, we estimated performance.

The accuracy, sensitivity, and specificity of these several ML algorithms were compared using three performance parameters, such as accuracy, sensitivity, and specificity.

The best results were achieved by the Random Forest algorithm, which had an accuracy of 97.92 percent, a sensitivity of 94.66 percent, and a specificity of 69.76 percent based on various performance parameters. Furthermore, Random Forest classification revealed a high level of discriminative ability. Along these lines, we can assume that the RF algorithm predicts nutritional status among Tanzania's under-five children moderately better than any of the other Machine Learning algorithms used in this study. However the Random Forest algorithm has the best prediction power when it comes to predicting childhood anemia (Khare et al., 2017). (Khan et al., 2019) identified several key traits. This could be because of the dataset they utilized. Finally, our findings suggest that when malnutrition prediction is a primary concern in Tanzania, random forest classification with random forest feature selection should be used.

5 Conclusions

We compared several machine learning algorithms for identifying whether a child is malnourished based on some risk variables in this study. The Random Forest algorithm outperformed the other algorithms in terms of classification accuracy for predicting malnutrition in Tanzania children. This study looks at the utility of Machine Learning calculations and mobile application as well as the importance of using common socio-demographic and health-related characteristics to predict nutritional status. Furthermore, our findings would be useful for subsequently identifying children at risk of malnutrition, providing policymakers and

medical service providers with a tool to make vital interventions and enhance care practices. As a result, a model based on the key risk factors could help prevent and control child malnutrition.

References

Alghamdi M, Al-Mallah M, Keteyian S, Brawner C, Ehrman J, Sakr S. 2017. Predicting diabetes mellitus using SMOTE and ensemble machine learning approach: The Henry Ford ExercIse Testing (FIT) project. PloS One, 12(7): e0179805

Anand P, Gupta R, Sharma A. 2020. Prediction of Anaemia among children using Machine Learning Algorithms. International Journal of Electronics Engineering, 11(2): 469-480

Browne C, Matteson DS, McBride L, Hu L, Liu Y, Sun Y, Wen J, Barrett CB. 2021. Multivariate random forest prediction of poverty and malnutrition prevalence. PloS One, 16(9): e0255519

Fenta HM, Zewotir T, Muluneh EK. 2021. A machine learning classifier approach for identifying the determinants of under-five child undernutrition in Ethiopian administrative zones. BMC Medical Informatics and Decision Making, 21(1): 1-12

Ion-Mărgineanu A, Kocevar G, Stamile C, Sima DM, Durand-Dubief F, Van Huffel S, Sappey-Marinier D. 2017. Machine learning approach for classifying multiple sclerosis courses by combining clinical data with lesion loads and magnetic resonance metabolic features. Frontiers in Neuroscience, 11: 398

Kaushik H, Singh D, Kaur M, Alshazly H, Zaguia A, Hamam H. 2021. Diabetic retinopathy diagnosis from fundus images using stacked generalization of deep models. IEEE Access, 9: 108276–108292

Khan JR, Chowdhury S, Islam H, Raheem E. 2019. Machine learning algorithms to predict the childhood anemia in Bangladesh. Journal of Data Science, 17(1): 195-218

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: 338-349. https://doi.org/10.1016/j.procs.2017.09.087

Kilicarslan S, Celik M, Sahin Ş. 2021. Hybrid models based on genetic algorithm and deep learning algorithms for nutritional Anemia disease classification. Biomedical Signal Processing and Control, 63: 102231

Lai H, Huang H, Keshavjee K, Guergachi A, Gao X. 2019. Predictive models for diabetes mellitus using machine learning techniques. BMC Endocrine Disorders, 19(1): 1-9

Tanzania Food and Nutrition Center, T. 2015. Tanzania National Nutrition Survey 2014. Tanzania

Tanzania National Nutrition Survey. 2018. Final Report. 2019. Tanzania Food and Nutrition Centre, National Bureau of Statistics, Tanzania

The World Health report. 2003. Shaping the Future. 2003. WHO, Switzerland

World Health Organization. 2021. The State of Food Security and Nutrition in the World 2021: Transforming Food Systems For Food Security, Improved Nutrition and Affordable Healthy Diets For All (Vol. 2021). Food & Agriculture Organization, Rome, Italy

Zhang WJ. 2010. Computational Ecology: Artificial Neural Networks and Their Applications. World Scientific, Singapore

Zhao Y, Healy BC, Rotstein D, Guttmann CR, Bakshi R, Weiner HL, Brodley CE, Chitnis T. 2017. Exploration of machine learning techniques in predicting multiple sclerosis disease course. PloS One, 12(4): e0174866

Zou Q, Qu K, Luo Y, Yin D, Ju Y, Tang H. 2018. Predicting Diabetes Mellitus With Machine Learning Techniques. Frontiers in Genetics, 9: 515. https://doi.org/10.3389/fgene.2018.00515