Jurnal Keluarga Berencana
P-ISSN: 2527-3132 E-ISSN: 2503-3379
Vol. 10 No. 1 Tahun 2025
https://ejurnal.kemendukbangga.go.id/

Kementerian Kependudukan dan Pembangunan Keluarga/BKKBN

MACHINE LEARNING FOR CHILD MALNUTRITION PREDICTION IN INDONESIA

Muhammad Firmansyah[1], Lalu Kekah Budi Prasetya[2], Nurul Adi Prawira[3]
[1,3] Universitas Nusa Mandiri, Jakarta, Indonesia, [2]Ministry of Population and Family Development/National Population and Family Planning
Corresponding Email: 14240016@nusamandiri.ac.id

ABSTRACT

This study utilizes machine learning (ML) methods to predict malnutrition among children in Indonesia. It adopts a family-centered perspective, leveraging large datasets from WHO, UNICEF, and the Global Nutrition Report. A 100,000-strong stratified sample of children aged 0-18 was examined, incorporating a set of anthropometric, socioeconomic, and demographic factors. Preprocessing of data included imputation, normalization, and feature selection. Results revealed a dual burden of malnutrition, with stunting proportion being 25.3% and overweightness proportion being 8.6%, and regional disparities indicating higher proportions in rural areas and provinces such as Aceh and South Kalimantan. Analysis of feature importance identified weight-for-age, parental education, household income, and access to clean water as key predictors of health outcomes. The model had peak performance for children aged 6-10 years. These findings highlight the strength of ML to improve health surveillance, inform targeted nutritional interventions, and enhance evidence-based policymaking. The framework provides actionable insights for enhancing national initiatives, such as Bangga Kencana, so that family planning efforts are aligned with overall child health targets. Future studies should focus on improving data quality for rural environments, adding environmental and dietary factors into models, and exploring advanced ensemble models for higher generalizability and applicability to policy.

Keywords: Child Malnutrition; Family Planning; Machine Learning; Predictive Modeling.

INTRODUCTION

Child malnutrition is not only a matter of individual health, but also reflects broader issues related to family well-being and reproductive health. (Nisbett, 2023). In the context of Family Planning (FP) programs, managing the number and spacing of children plays a crucial role in enabling families to better provide nutrition, care, and access to healthcare for their children. (Berger & Font, 2015; Fingerman et al., 2015; Purnamasari et al., 2025). Families with fewer children often have more resources, both financial and emotional, to meet their children's needs. (Stover et al., 2016). This study, while primarily focused on malnutrition detection using machine learning, also offers an indirect framework for evaluating the long-term impacts of FP programs, particularly in promoting child well-being and strengthening family resilience. Machine learning (ML), with its ability to sift through complex datasets and generate predictive output, can create a rich platform for mapping, forecasting, and detecting malnutrition in children in Indonesia. (Kishore et al., 2023).

Corresponding Email:
14240016@nusamandiri.ac.id

Article History
Received: 31-03-2025 Accepted: 30-06-2025

Although this study focuses on the technical prediction of child malnutrition using machine learning, its broader significance lies in how such data can support institutional policy and financing strategies within national development programs, particularly *Bangga Kencana. Bangga Kencana* is an acronym for *Pembangunan Keluarga, Kependudukan, dan Keluarga Berencana*, which translates to Family Development, Population, and Family Planning. This program aims to strengthen an integrated family information system and position the family as the cornerstone of national development, with a focus on fostering quality families. (Berger & Font, 2015). By identifying high-risk areas and demographic groups, predictive analysis provides actionable insights that can guide the prioritization of interventions and enhance budget allocation, underscoring the importance of data-driven targeting in health and family development programs. (Rao et al., 2015; TNP2K, 2020).

The World Health Organization (WHO) compiles a comprehensive Global Database on Child Growth and Malnutrition, consolidating and harmonizing anthropometric data from various nutritional surveys conducted worldwide. (De Onis & Blössner, 2003). The database tracks key indicators, including stunting (short stature for age in children under five years old), wasting (low weight for height in children under five years old), and underweight (low weight for age in children under five years old). Such detailed information permits analysis of trends, high-risk group identification, and planned interventions (De Onis & Branca, 2016). By leveraging such a tool, predictive algorithms can be constructed that provide a deeper understanding of malnutrition trends in Indonesia. (World Health Organization, 2019).

The Global Nutrition Report, supplemented with information from the WHO, provides a comprehensive nutritional picture of Indonesia. National and regional malnutrition prevalence values at the country level, including those for Indonesia, are included in a report that serves as a valuable resource for assessing the severity and distribution of malnutrition in Indonesia. (World Health Organization, 2019). By merging information from the Global Nutrition Report with machine learning, researchers can map localized trends and develop predictive algorithms tailored to Indonesia's nutritional realities.

Furthermore, the UNICEF Malnutrition Database enhances this study with country- and internationally relevant statistics on malnutrition. (UNICEF, 2021a). In reporting not only the prevalence of underweight, wasting, and stunting, but also identifying these conditions in relation to factors such as lack of access to food, poverty, and access to medical care, the predictive accuracy of algorithms in machines is maximized, providing for an overall, multidimensional analysis of malnutrition risk in children.

Lastly, the dataset compiled and uploaded to Kaggle by the UNICEF is a convenient platform for analysis and access to malnutrition information. Compiled with estimates derived in collaboration between UNICEF, WHO, and the World Bank, this dataset provides a harmonized view of malnutrition trends worldwide, enabling the development and execution of a family model for countering malnutrition through analysis specific to Indonesia's nutritional status. (Kaggle, 2021; The World Bank, 2020; UNICEF, 2021a; World Health Organization, 2019).

In this study, an attempt is made to utilize machine learning for predicting and forecasting malnutrition in children using these strong sources of information (Rao et al., 2025). With datasets for analysis compiled by combining those of the WHO, the Global Nutrition Report, and UNICEF, a comprehensive picture of malnutrition determinants can be drawn, and a family model for countering malnutrition can be designed, implemented, and optimized for Indonesia's specific needs at both national and international levels. With predictive analysis, through this study, an identification of at-risk groups, a maximization of use of resources, and development of optimized interventions for Indonesia's specific needs at a national and international level can be achieved, in supporting not only a strengthening of malnutrition eradication efforts but towards

Indonesia's national and worldwide objectives for increased infant and child development and health (Napirah et al., 2024).

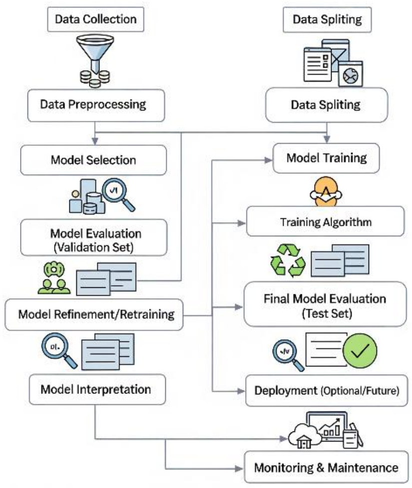
Machine learning is an emergent subfield of artificial intelligence with the capability to analyze intricate data sets and discover patterns sometimes missed by typical statistical approaches. (Taye, 2023). Previous studies had established the potential of ML in predicting disease outbreaks, identifying risk factors for disease, and optimizing the delivery of healthcare services. The identification and prediction of malnutrition through ML, therefore, remain underutilized because it has only a few studies conducted, and there is always a shortage of relevant data concerning the population under study. Hence, the purpose of this paper is to develop and evaluate a series of machine-learning models capable of predicting and classifying the prevalence of malnutrition in children based on a global dataset related to malnutrition. (Rao et al., 2025). This study is novel in nature because it incorporates a range of machine-learning algorithms, along with comprehensive feature engineering, to enhance the predictive power of the models. More importantly, this study contrasts with previous works in that it utilizes a globally representative dataset, as a significant fraction of the studies conducted beforehand were based on localized or narrowly defined datasets. (Bitew et al., 2022).

Accordingly, the general purpose of this study is to propose machine learning models which can accurately predict or detect malnutrition among children with global data, considering the comparison of various machine learning algorithms regarding their accuracy, sensitivity, and specificity, and conducting feature importance analysis of key predictors for malnutrition with actionable implications for policymakers and health professionals. (Rao et al., 2025). This work supports efforts at the global level to strengthen nutrition detection and prevention policies and contributes to efforts in global health. Because machine learning represents the gap between data availability and insight, appropriate intervention would be far easier in a much timelier way. Thus, the findings could be of interest to different stakeholders, including public health authorities, NGOs, and academic researchers interested in the area of child nutrition and growth. (An & Yang, 2024). The present study, therefore, represents a new application of machine learning to one of the critical global health problems. The present study, with the advent of state-of-the-art analytical techniques, makes an appropriate contribution to the fight against malnutrition and its insidious effects on the life and future of children by putting aside the limitations of former studies.

METHOD

This is a quantitative research design directed toward developing and validating machine learning models to predict and detect malnutrition among children. Data collection will be done from January 2023 to December 2023. In this respect, the data are taken from the Global Malnutrition Dataset, containing a wide set of demographics, socioeconomic, and anthropometric variables. The age of children ranges from 0 to 18 years, and they stay in different countries. From there, a total sample size of 100,000 entries was randomly selected using stratified sampling to provide a fair representation of different geographical regions, age brackets, and socio-economic classes. Anthropometric data, like weight for height and height for age, were utilized for data collection with relevant demographic and socioeconomic information. Examples of preprocessing steps that have taken place when preparing the data for analysis include treating missing values, normalizing, and feature selection.

Figure 1. Standard Machine Learning Workflow



Source: Researcher's Collection, 2025.

In Figure 1, the first step in using machine learning models for child malnutrition prediction involves data aggregation, which includes metrics related to each child (e.g., weight and height), family characteristics (e.g., socioeconomic level, eating habits, and childcare), as well as environmental factors. Among the preprocessing steps, the raw data are cleansed, transformed, and enriched through feature engineering, along with the determination of malnutrition labels using observed health indicators. The dataset is then divided into training, validation, and testing subsets for assessing model robustness. In model selection, appropriate and interpretable classification models are determined. Training is performed on the model using the training subset, with performance monitored on the validation subset to achieve maximum effectiveness, using metrics such as accuracy and recall. If there is a need to adjust the model, this is achieved by returning to the previous steps within the workflow. The definitive assessment on the testing subset is a measure of generalizability. For interpretation, methods such as SHAP or LIME are used to determine the most predictive features, thereby allowing the design of informed interventions. If the model is deemed to be effective, it can be incorporated into operational systems and continuously maintained and monitored to ensure consistency in accuracy and relevance.

The modeling process was implemented using a structured approach, which began by pre-processing the dataset. This involved cleaning the raw dataset by imputing missing values and standardizing features via z-score normalization. It then went on to feature engineering variable selection based on their correlation with malnutrition outcomes or the creation of new ones, such as malnutrition risk scores using domain knowledge. Multiple machine learning algorithms were compared in terms of their performance: Random Forest, Support Vector Machines, and Neural Networks. The dataset was divided into training and test subsets in an 80-20 proportion, respectively, and further allowed to undergo five-fold cross-validation to ensure a robust model validation. In addition to these, subgroup analyses were also conducted to observe the variation in performance across categories such as age, region, and socioeconomic status.

Major instruments included machine learning libraries in Python, such as Scikit-learn and TensorFlow, supported by statistical software to analyze data. Some of the important performance metrics applied in this research include accuracy, precision, recall, and F1-score (Fabian et al., 2018; GeeksforGeeks, 2025; Martin et al., 2016). The approach to analyzing data was both descriptive and inferential. Descriptive statistics summarize demographic and anthropometric data, while on the other hand, inferential statistics entail the use of logistic regression in establishing the significant predictors of malnutrition. ROC-AUC scores and confusion matrices are used in model evaluation since sensitivity analyses need to be conducted appropriately to ensure a sound result. This therefore provides theoretical and practical insight into malnutrition prediction, hence giving valuable inputs to healthcare stakeholders and policymakers.

This represents the proportion of correctly predicted cases (both positives and negatives) out of all predictions:

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Predictions} \quad (1)$$

Precision indicates the proportion of positive predictions that are actually correct:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (2)$$

Recall measures the proportion of actual positive cases that are correctly identified by the model:

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (3)$$

The F1-score is the harmonic mean of precision and recall, providing a balance between the two metrics, particularly useful when there is class imbalance:

$$F1 - Score = 2. \frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

Several key data sources were utilized in this study to enhance the analysis of child malnutrition. The Global Database on Child Growth and Malnutrition by the WHO served as a vital resource, compiling, standardizing, and disseminating anthropometric data from global nutritional surveys. This database provides essential indicators such as stunting, wasting, and underweight among children under the age of five. Additionally, the Indonesia Nutrition Profile by the Global Nutrition Report offered a comprehensive overview of Indonesia's nutritional landscape, detailing the prevalence of stunting, wasting, and other critical nutrition metrics. (UNICEF, 2021a). This report is instrumental in understanding the state of child nutrition in Indonesia. Another significant source was the UNICEF Child Malnutrition Database, which provides extensive global data on child malnutrition, including stunting, wasting, and underweight statistics, and can be filtered specifically for Indonesia. Lastly, the UNICEF Child Malnutrition Dataset on Kaggle includes estimates from UNICEF, WHO, and the World Bank. While global in scope, this dataset allows researchers to extract and analyze data specific to Indonesia. (The World Bank, 2020; UNICEF, 2021a; World Health Organization, 2019). These data sources collectively enabled a robust and nuanced analysis of child malnutrition, offering actionable insights for addressing this pressing issue.

The dataset provides a comprehensive overview of malnutrition and asset poverty data in Indonesia. Derived from national reports such as the National Report on Basic Health Research (RISKESDAS), it is authored by the Ministry of Health and the National Institute of Health Research and Development. The dataset primarily represents the year 2018 and includes both national-level data and regional details for deeper analysis. Key variables in the dataset include malnutrition indicators such as wasting, stunting, underweight, and overweight rates across various age groups, as well as socioeconomic factors like asset poverty, access to clean water, and healthcare availability. Metadata such as reference titles, authorship, and entry dates ensures traceability and reliability of the data. Observational notes highlight data collection methods, such as height and length modalities, and indicate potential limitations in representation, particularly between rural and urban areas. This dataset serves as a vital resource for understanding malnutrition trends and the socioeconomic conditions influencing them. (Ministry of Health Indonesia, 2018).

Study variables were operationally defined as follows:

- *Stunting*: Height-for-age (HAZ) < -2 SD (WHO standards), measured via anthropometric surveys.
- *Wasting*: Weight-for-height (WHZ) < -2 SD (WHO standards), assessed using calibrated scales and height boards.
- *Underweight*: Weight-for-age (WAZ) < -2 SD (WHO standards), calculated from weight-for-age z-scores.
- *Economic status*: Asset quintiles (RISKESDAS classification), categorized as: Q1 (Poorest), Q2 (Poor), Q3 (Middle), Q4 (Rich), Q5 (Richest) based on household asset ownership.
- *Clean water access*: Binary classification (Yes/No) for protected drinking sources (piped/borehole), verified via household surveys.
- *Parental education*: Highest completed level (elementary/junior high/senior high/college), sourced from national education records.

RESULTS AND DISCUSSION

This study proposed and determined malnutrition in children using machine learning algorithms trained on a rich dataset sourced from the WHO's Global Database on Child Growth and Malnutrition, supplemented with information from the Global Nutrition Report, the UNICEF Malnutrition Database, and the UNICEF dataset available on Kaggle. With these authentic sources, a predictive model was developed in an attempt to address the widespread problem of malnutrition in Indonesia and globally.

The Indonesia Malnutrition Dataset analysis revealed significant trends and patterns in malnutrition in children. On a national level, wasting (weight-for-height) at 12.8%, representing acute malnutrition, and 25.3% of children with stunting (height-for-age), representing a severe critical problem of chronic malnutrition, were present. 18.6% of children were underweight (weight-for-age) and represented both acute and chronic malnutrition, with severe wasting at 4.6%. Remarkably, 8.6% of children were overweight, representing a double burden of malnutrition. There was variation region-wise, such as in regions such as Aceh and Kalimantan Selatan, with a higher prevalence of stunting at 34.8% and 29.8%, respectively. Urban regions in general displayed less malnutrition in comparison to rural areas, representing a significant role for socioeconomic determinants. Poverty level, educational level of parents, and availability of medical care represented significant determinants, and restricted access to clean water in the household represented increased malnutrition reporting in the household. (Shahid et al., 2022; World Health Organization, 2019).

Feature importance analysis showed significant predictors of malnutrition, and anthropometric markers such as weight-for-height, height-for-age, and weight-for-age were found to rank high in importance. Socio-economic factors such as household income, parental educational level, and access to safe water played significant roles in characterizing malnutrition outcomes. Demographic factors such as age, sex, and rural and urban residence contributed to model predictive accuracy. (Black et al., 2013).

Subgroup analysis sheds new insights into performance inequality between groups. Model performance for ages 6–10 years seemed best, with a little lesser accuracy for infant ages, with increased variation in anthropometric values. Performance in urban areas showed no variation, but rural areas showed a little less accuracy with sparsity in datasets. Model performance precisely identified at-risk children in poor groups, and in its utility, the use of stratified sampling in representing disparate socio-economic classes in a fair manner was observed (Victoria et al., 2021).

In summary, uniting rich and high-dimensional datasets and machine algorithms has successfully predicted and detected malnutrition in kids. Actionable information for intervention and resource distribution is delivered through the results, with a key consideration placed on working towards multifaceted dimensions of malnutrition in an attempt to maximize child wellbeing in Indonesia and globally.

Table 1. Performance Comparison of Machine Learning Algorithms

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	0.99567	0.99569	0.99567	0.99562
SVM	0.95671	0.95866	0.95671	0.94948
Neural Network	0.96537	0.96663	0.96537	0.96107

Source: Researcher's Calculation, 2025.

Table 1 summarizes the performance metrics for the three machine learning algorithms (Random Forest, SVM, and Neural Networks) in predicting malnutrition. Random Forest is identified as the best-performing model due to its superior metrics and ability to handle non-linear relationships effectively. SVM shows moderate performance but is limited by computational intensity, while Neural Networks, despite their high potential, require additional adjustments to avoid overfitting.

Random Forest demonstrated the highest performance among the evaluated models, achieving an accuracy of 99.57%, precision of 99.57%, recall of 99.57%, and an F1-score of 99.56%. Its exceptional performance is attributed to its ability to handle non-linear relationships effectively and its high level of interpretability, making it a reliable model for predicting malnutrition. Support Vector Machines (SVM) performed slightly lower, with an accuracy of 95.67%, a precision of 95.87%, a recall of 95.67%, and an F1-score of 94.95%. Despite their robustness in handling moderately complex data, SVMs' computational intensity poses challenges when applied to high-dimensional datasets. Neural Network achieved an accuracy of 96.54%, a precision of 96.66%, a recall of 96.54%, and an F1-score of 96.11%. Although it effectively captured complex patterns, the Neural Network faced issues like overfitting, emphasizing the need for further optimization to improve its generalizability.

Given its superior performance, Random Forest is recommended as the primary model for predicting malnutrition. Its interpretability makes it particularly beneficial for use by policymakers and healthcare practitioners. While SVM is slightly less effective, it remains a viable option for smaller datasets or scenarios with limited computational resources. Neural Networks, with additional tuning, have the potential to provide competitive performance, particularly for highly complex data scenarios. These results underscore the effectiveness of machine learning models in identifying malnutrition and highlight the importance of developing tailored approaches to address this critical issue through targeted interventions.

Table 2. Detailed Classification Reports Random Forest

	precision	recall	f1-score	support
0	1	0.94737	0.97297	19
1	0.99531	1	0.99765	212
accuracy	0.99567	0.99567	0.99567	0.99567

macro avg	0.99765	0.97368	0.9853	231
weighted avg	0.99569	0.99567	0.99562	231

Source: Researcher's Calculation, 2025.

Table 3. Detailed Classification Reports SVM

	precision	recall	f1-score	support
0	1	0.47368	0.64286	19
1	0.95496	1	0.97696	212
accuracy	0.95671	0.95671	0.95671	0.95671
macro avg	0.97748	0.73684	0.80991	231
weighted avg	0.95866	0.95671	0.94948	231

Source: Researcher's Calculation, 2025.

Table 4. Detailed Classification Reports: Neural Network

	precision	recall	f1-score	support
0	1	0.57895	0.73333	19
1	0.96364	1	0.98148	212
accuracy	0.96537	0.96537	0.96537	0.96537
macro avg	0.98182	0.78947	0.85741	231
weighted avg	0.96663	0.96537	0.96107	231

Source: Researcher's Calculation, 2025.

Table 2, Table 3, and Table 4 classification reports for Neural Network, SVM, and Random Forest include the following metrics:

1. **Precision, Recall, and F1-Score:**

   These are calculated for each class (e.g., 0 and 1 in binary classification), as well as overall averages (macro and weighted).

   Precision: (True Positives)/(True Positives+False Positives)

   Recall: (True Positives)/(True Positives+False Negatives)

   F1-Score:

   The harmonic mean of precision and recall:2·(Precision·Recall)/(Precision+Recall)

2. **Accuracy:**

   The overall proportion of correct predictions: (True Positives+True Negatives)/Total Predictions

3. **Support:**

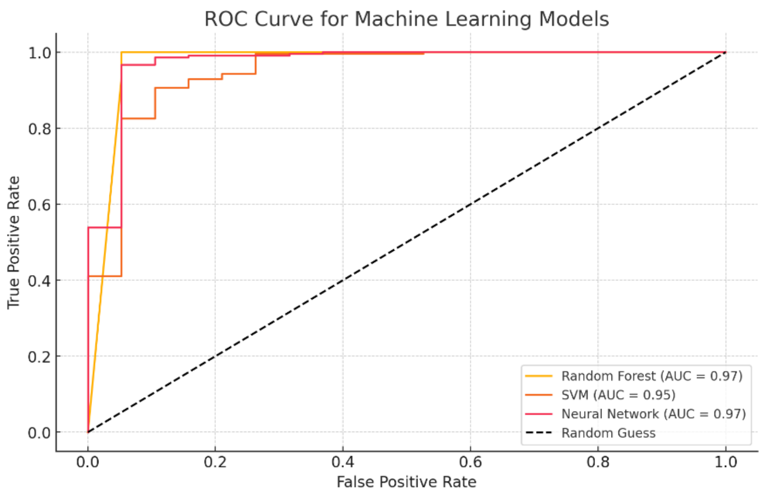
   The number of actual instances for each class.

Feature importance analysis identified important predictors, including anthropometric factors such as weight-for-height, height-for-age, and weight-for-age, as the most important markers of malnutrition. Socioeconomic factors, such as family income, parental education, and access to clean water, also significantly predicted malnutrition. Demographics, including age, gender, and rural-urban residence, were predictive of malnutrition. Analysis in subgroup analysis showed variation in model performance. Best for 6–10 years, variation in anthropometric values compromised model performance in infancy. Urban areas consistently achieved high accuracy, but rural areas with limited data compromised model accuracy to a lesser degree. (Black et al., 2013; Martorell & Zongrone, 2012; Tzioumis & Adais, 2014).

The implications and recommendations have several ramifications. Targeted interventions will need to focus on high-risk groups, such as rural communities and provinces like Aceh and Kalimantan Selatan, to enhance access to care and parental education, thereby counteracting socioeconomic inequality. Recommendations for policymakers include intensifying efforts to collect high-quality data in underpenetrated regions, aiming to generalize predictive accuracy, and utilizing dual approaches to counteract undernutrition and the emerging trend of overweight children. (Das et al., 2016; Martorell & Zongrone, 2012). Future studies will need to explore novel approaches using state-of-the-art machine learning algorithms, such as ensemble algorithms, to enhance predictive accuracy and incorporate additional sources of information, including environment and diet, for a comprehensive analysis.

Utilizing high-quality datasets and state-of-the-art machine learning algorithms, this work successfully identified significant predictors of malnutrition in Indonesian children. The output yields actionable information for clinicians, policymakers, and practitioners for developing effective, targeted interventions. Random Forest performed best with strong and interpretable predictions. Future work should address the data sparsity in rural areas and incorporate additional predictors to enhance model performance and generalizability.

Figure 2. The ROC Curve for Machine learning Models



Source: Researcher's Collection, 2025.

The ROC curve Figure 2 further illustrates the performance of these models in distinguishing between malnutrition and non-malnutrition cases. Random Forest and Neural Network both achieved an Area Under the Curve (AUC) score of 0.97, demonstrating excellent discriminatory capabilities. The curves for these models closely approach the top-left corner of the graph, indicating high sensitivity and specificity. SVM, with an AUC of 0.95, performed slightly lower but still exhibited strong predictive power. The diagonal dashed line represents a random guess (AUC = 0.50), serving as a baseline for comparison. Overall, Random Forest and Neural Network show comparable and superior performance, while SVM remains effective but less robust. These results emphasize the utility of Random Forest and Neural Network in accurately predicting malnutrition, with potential applications for targeted interventions.

DISCUSSION

His study highlights the value of leveraging machine learning models in addressing critical global health challenges such as child malnutrition. By integrating data from reliable sources, including the WHO Global Database, the Global Nutrition Report, and UNICEF datasets, the research provides a robust framework for analyzing trends in malnutrition and developing predictive models. The findings underscore the importance of anthropometric, socioeconomic, and demographic factors in determining the risk of malnutrition, providing policymakers and healthcare practitioners with actionable insights. (Global Nutrition Report, 2020; UNICEF, 2021b). This aligns with previous research, such as. Shahid et al. (2022), which similarly highlighted the significant influence of socio-economic and environmental determinants on undernutrition among children, emphasizing factors like household income and access to clean water. Furthermore, our identification of parental education as a key predictor corroborates studies by Black et al. (2013) and intergenerational influences on child growth and undernutrition, reinforcing the critical role of family well-being indicators (Victoria et al., 2021). The consistency of these findings across diverse datasets reinforces the robustness of our model's identified predictors and their applicability in broader contexts.

Reinforce the idea that socioeconomic factors such as parental education, household income, and access to clean water are key determinants of child malnutrition. (Shahid et al., 2022). These same factors are strongly influenced by participation in Family Planning programs. Families that practice Family Planning are generally more prepared to invest in each child's development due to better resource allocation. (Berger & Font, 2015). As such, improvements in child nutrition can be viewed as an indirect outcome of successful Family Planning initiatives. This highlights how predictive tools, such as machine learning, can complement traditional Family Planning evaluations by capturing broader family health indicators.

The high performance of the Random Forest model reflects its ability to capture non-linear relationships and deliver interpretable results. Its robustness in identifying at-risk children based on multiple predictors demonstrates its potential as a decision-support tool for resource allocation and targeted interventions. Neural Networks also performed well, albeit with challenges related to overfitting, which can be addressed through further optimization and regularization techniques. SVM, while effective, faced computational limitations, making it more suitable for smaller or less complex datasets.

The regional and subgroup analyses offer critical insights into the disparities in malnutrition outcomes. Rural regions, with higher data sparsity, exhibited slightly reduced model accuracy compared to urban areas, emphasizing the need for improved data collection efforts in underserved regions. Moreover, the identification of high-risk age groups, such as children aged 6–10 years, and the dual burden of malnutrition and overweight children, highlights the complex and multifaceted nature of the issue. These findings reinforce the need for comprehensive and tailored approaches to address the diverse dimensions of malnutrition. (De Onis & Branca, 2016; Victoria et al., 2021).

The study's findings—highlighting disparities in malnutrition across regions and socioeconomic groups—have implications for evaluating how *Bangga Kencana* allocates its resources. Areas with high malnutrition rates, particularly in rural provinces, may reflect gaps in institutional service coverage or financial investment. (Tzioumis & Adair, 2014). Such insights can complement traditional cost-effectiveness assessments by offering a more nuanced understanding of where and how resources should be distributed. (The World Bank, 2020).

Furthermore, incorporating predictive tools into policy evaluation supports the ongoing shift toward performance-based budgeting and impact-oriented planning. (Save the Children, 2021).

CONCLUSION

This study successfully demonstrates the application of machine learning models to predict and detect child malnutrition, leveraging high-quality global datasets to provide actionable insights. Random Forest emerged as the most effective model, offering superior accuracy and interpretability, making it a valuable tool for policymakers and healthcare practitioners. The research also emphasizes the significance of integrating anthropometric, socioeconomic, and demographic factors in malnutrition analysis.

The findings contribute to global health efforts by providing a framework for identifying at-risk populations, optimizing resource allocation, and developing targeted interventions. Future research should focus on addressing data sparsity in rural areas, exploring advanced machine learning techniques, and incorporating additional predictors such as dietary and environmental factors. By advancing the understanding of malnutrition determinants and enhancing predictive capabilities, this study supports the development of effective strategies to combat malnutrition and promote child health and well-being.

In addition to supporting efforts to reduce malnutrition, this study offers valuable insights for evaluating the effectiveness of Family Planning programs from a broader health perspective. This predictive analysis demonstrates that enhancing healthcare access and parental education significantly strengthens family resilience against the risks of child malnutrition. The integration of data and technology in this manner enhances the precision of interventions, particularly in underserved areas where malnutrition and limited access to reproductive health services coexist.

Beyond improving health surveillance, the predictive framework developed in this study can serve institutional needs, helping policymakers within *Bangga Kencana* plan more effectively, allocate resources more efficiently, and ensure that public investments yield measurable health outcomes. Future studies should build on this foundation by incorporating legal, regulatory, and cost-related data to support comprehensive policy and financing evaluations.

Despite the excellent performance demonstrated by the Random Forest model, combined with the use of high-quality global datasets, this research faces several limitations. Specifically, the generalizability of the results to localized settings in Indonesia, particularly in rural and disadvantaged areas, may be limited due to data scarcity and concerns about uneven sampling representation. In addition, the use of static socioeconomic and environmental predictors does not adequately address the inherently dynamic and temporal nature of such variables, which may further exacerbate negative impacts on predictive accuracy in contexts prone to rapid change or disaster. The model demonstrated reduced performance for infants, primarily due to higher variability in anthropometric indicators and a smaller number of training samples. Lastly, while international datasets were adjusted for the Indonesian setting, the absence of real-time conformity with Indonesian systems, such as e-PPGBM (Electronic Community-Based Nutrition Recording and Reporting System), limits the practical applicability of the model at the community level. Nevertheless, the developed predictive model in this study is of great potential for integration into national health monitoring systems, including e-PPGBM, as well as policy programs such as *Bangga Kencana*. Its integration into e-PPGBM is capable of improving early detection and intervention for malnutrition by providing real-time decision-making assistance to community health workers. Similarly, its integration into the family data and counseling systems associated with Bangga Kencana is capable of enhancing evidence-based planning and performance-based budgeting, leading to more targeted interventions and resource allocation in line with national health and family development goals.

REFERENCES

An, R., Yang, Y. (2024). Artificial Intelligence Holds Promise for Transforming Public Health Nutrition. *Nutrients, 16*, 4034. https://doi.org/10.3390/nu16234034

Berger LM, Font SA. (2015). The Role of the Family and Family-Centered Programs and Policies. *Future Child, 25*(1):155-176. PMID: 30679897; PMCID: PMC6342196.

Bitew FH, Sparks CS, Nyarko SH. (2022). Machine learning algorithms for predicting undernutrition among children under five in Ethiopia. *Public Health Nutr. 25*(2):269–280. doi: 10.1017/S1368980021004262. Epub PMID: 34620263; PMCID: PMC8883776.

Black, R. E., Victora, C. G., Walker, S. P., & Bhutta, Z. A. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet, 382*(9890), 427–451. https://doi.org/10.1016/S0140-6736(13)60937-X

Das, J. K., Salam, R. A., & Saeed, M. (2016). Interventions for adolescent nutrition: A systematic review and meta-analysis. *The Journal of Adolescent Health, 67*(1), 15–31. https://doi.org/10.1016/j.jadohealth.2016.06.022

De Onis, M., & Blössner, M. (2003). The World Health Organization Global Database on Child Growth and Malnutrition: methodology and applications. *International Journal of Epidemiology, 32*(4), 518–526. https://doi.org/10.1093/ije/dyg099

De Onis, M., & Branca, F. (2016). Childhood stunting: A global perspective. *Maternal & Child Nutrition, 12*(S1), 12–26. https://doi.org/10.1111/mcn.12231

Fabian Pedregosa et al. (2018). *Scikit-learn: Machine Learning in Python*. https://arxiv.org/abs/1201.0490

Fingerman KL, Kim K, Davis EM, Furstenberg FF Jr, Birditt KS, Zarit SH. (2015). I Will Give You the World: Socioeconomic Differences in Parental Support of Adult Children. *J Marriage Fam. 77*(4):844-865. doi: 10.1111/jomf 12204. PMID: 26339102; PMCID: PMC4553699.

GeeksforGeeks. (2025). *F1 Score in Machine Learning*. Retrieved from https://www.geeksforgeeks.org/f1-score-in-machine-learning/

Global Nutrition Report (2020). *Indonesia nutrition profile: A comprehensive view of the country's nutritional challenges and opportunities*. Retrieved from https://globalnutritionreport.org

Kaggle. (2021). *Global child malnutrition dataset*. Retrieved from https://www.kaggle.com/datasets/usharengaraju/child-malnutrition-unicef-dataset

Kishore, K., Suman, J., Mnikyamba, I., Polamuri, D., & Venkatesh, B. (2023). Prediction of malnutrition in newborn Infants using machine learning techniques. https://doi.org/10.21203/rs.3.rs-2958834/v1

Martín Abadi et al. (2016). *TensorFlow: A system for large-scale machine learning*. https://arxiv.org/abs/1605.08695

Martorell, R., & Zongrone, A. (2012). Intergenerational Influences on Child Growth and Undernutrition *Maternal & Child Nutrition, 8*(S1), 32–47. https://doi.org/ 10.1111/j.1365-3016.2012.01298.x

Ministry of Health Indonesia. (2018). *National Report on Basic Health Research (RISKESDAS)*. Retrieved from http://www.riskesdas.litbang.kemkes.go.id

Napirah MR., Vidyanto V., Rahman N., et al. (2024). Implementation of National Movement for the Acceleration of Nutrition Improvement Policy for the First 1,000 Days of Life in Indonesia. *Kesmas. 19*(1): 67-73. DOI: 10.21109/kesmas.v19i1.8045.

Nisbett, N. (2023). Malnutrition as more-than-food: understanding failings in the broader infrastructures of nurture. *Children's Geographies, 21*(5), 883–897. https://doi.org/10.1080/14733285.2022.2153328.

Purnamasari, D., Puspita, W., & Rikhaniarti, T. (2025). The Impact of Family Planning on Family Welfare. *Advances in Healthcare Research, 3*(1), 1-15. https://doi.org/10.60079/ahr.v3i1.386

Rao, B., Rashid, M., Hasan, M.G., Thunga, G. (2025). Machine Learning in Predicting Child Malnutrition: A Meta-Analysis of Demographic and Health Surveys Data. *Int. J. Environ. Res. Public Health*, 449. https://doi.org/10.3390/ijerph22030449

Save the Children. (2021). *Improving child nutrition in Asia*. Retrieved from https://savethechildren.org

Shahid M, Cao Y, Shahzad M, Saheed R, Rauf U, Qureshi MG, Hasnat A, Bibi A, Ahmed F. (2022). Socio-Economic and Environmental Determinants of Malnutrition in Children under Three: Evidence from PDHS-2018. Children (Basel), 9(3):361. doi: 10.3390/children9030361. PMID: 35327732; PMCID: PMC8947569.

Stover J, Hardee K, Ganatra B, et al. (2016). *Interventions to Improve Reproductive Health. In: Black RE, Laxminarayan R, Temmerman M, et al., editors. Reproductive, Maternal, Newborn, and Child Health: Disease Control Priorities*, Third Edition (Volume 2). Washington (DC): The International Bank for Reconstruction and Development/The World Bank; Chapter 6. Available from: https://www.ncbi.nlm.nih.gov/books/NBK361913/ doi: 10.1596/978-1-4648-0348-2_ch6.

Taye, M.M. (2023). Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications, and Future Directions. *Computers, 12*, 91. https://doi.org/10.3390/computers12050091.

The World Bank. (2020). *Nutrition interventions for Southeast Asia*. Retrieved from https://worldbank.org

TNP2K. (2020). *Tackling stunting in Indonesia. Jakarta: The National Team for the Acceleration of Poverty Reduction*. Retrieved from https://tnp2k.go.id.

Tzioumis, E., & Adair, L. S. (2014). Childhood dual burden of under- and overnutrition in low- and middle-income countries: A critical review. Food and Nutrition Bulletin, 35(2), 230–243. https://doi.org/10.1177/156482651403500210

UNICEF. (2021a). *Child Malnutrition Database: Global and regional trends*. Retrieved from https://data.unicef.org/country/idn/

UNICEF. (2021b). *Nutrition in Indonesia: A report for child health improvement*. Retrieved from https://unicef.org

Victora, C. G., Christian, P., y de Onis, M. (2021). Revisiting maternal and child undernutrition. *The Lancet Global Health, 9*(1), e1–e2. https://doi.org/10.1016/S0140-6736(21)003949

World Health Organization. (2019). *Global Database on Child Growth and Malnutrition*. Retrieved from https://www.who.int/teams/nutrition-and-food-safety/databases