

K-Means and Fully Connected Neural Network for Child Nutritional Status Classification

Rismayanti^{1)*}, Sumi Khairani²⁾

¹⁾ Informatics Engineering Study Program, Faculty of Engineering and Computer Science, Universitas Harapan Medan, Medan, Indonesia, ²⁾ Informatics Management Study Program, Faculty of Engineering and Computer Science, Universitas Harapan Medan, Medan, Indonesia

¹⁾risma.stth@gmail.com, ²⁾sumibintisyafullah@gmail.com

Submitted : June 16, 2026 | **Accepted** : July 1, 2026 | **Published** : July 9, 2026

Abstract: Stunting remains a persistent child nutrition problem because delayed growth is closely related to long-term health, cognitive, and productivity risks. Manual interpretation of anthropometric measurements using the World Health Organization Z-score standard is clinically valid, yet it becomes inefficient and error-prone when routine records are processed in large numbers. This study develops a child nutritional status classification model by combining K-Means clustering and a fully connected neural network for early identification of stunting, underweight, and wasting. The dataset consisted of toddler anthropometric records from 2021-2024 with sex, age, body weight, and body height attributes. The data were cleaned, standardized, transformed into Z-score indicators, and grouped into 27 clusters representing possible combinations of nutritional status profiles. Cluster membership was then used with Zlen, Zwei, and Zwfl features in a multi-head fully connected neural network. Evaluation on 82 held-out samples showed accuracy values of 91.46% for stunting, 93.90% for underweight, and 98.78% for wasting. Weighted precision, recall, and F1-score were consistently high across the three outputs, while the training curves indicated stable learning without strong overfitting. The proposed hybrid model improves the reliability of child nutrition classification and can support a web-based decision support system for data-driven nutritional screening and intervention planning.

Keywords: Child nutrition; Fully connected neural network; K-Means clustering; Stunting; Z-score;

INTRODUCTION

Child malnutrition remains a major public health concern because it reflects not only inadequate dietary intake, but also the cumulative influence of infection, caregiving practices, sanitation, and socioeconomic conditions. The UNICEF, World Health Organization, and World Bank Joint Child Malnutrition Estimates reported that 150.2 million children under five years of age were stunted in 2024, while wasting and overweight also remained important global concerns (UNICEF, 2025). In Indonesia, stunting is still treated as a strategic development issue because it can affect human capital quality across generations (Lameky, 2024; Milwan & Sunarya, 2023).

Stunting is commonly interpreted through height-for-age Z-score, in which a child whose height is below minus two standard deviations from the World Health Organization standard is categorized as stunted. The same anthropometric framework also supports the identification of underweight through weight-for-age Z-score and wasting through weight-for-height or weight-for-length Z-score (Onis, 2006). These three indicators are closely related but represent different nutritional conditions. Stunting generally reflects chronic growth failure, underweight indicates low weight relative to age, and wasting reflects acute nutritional deficit relative to height or length.

Routine nutritional screening generally begins with measuring age, sex, body weight, and body height. These measurements are then converted into Z-score indicators using reference values (Nursanti et al., 2023; Yang et al., 2022). Although this method is medically accepted, its use in large-scale administrative or community datasets requires careful calculation and consistent interpretation. Manual processing may increase the risk of data-entry errors, inconsistent categorization, and delayed reporting. These limitations create a need for computational models

*name of corresponding author



that can transform anthropometric records into structured decision support outputs while maintaining compatibility with the World Health Organization standard.

Machine learning has been increasingly used to support nutritional classification and stunting prediction. Prior studies have explored decision tree, support vector machine, random forest, K-nearest neighbor, and clustering models for stunting-related tasks (Hasdyna et al., 2024; Kirk et al., 2022; Theodore Armand et al., 2024). These models can be effective, yet many studies treat a single nutritional indicator as an isolated classification target. In practice, stunting, underweight, and wasting are related manifestations of child growth and nutrition. A model that learns these outputs together may capture shared patterns better than a model that handles each task independently.

In addition to conventional machine learning, deep learning provides an alternative approach for learning non-linear relationships in structured numerical data (Ahmed et al., 2023; Liao et al., 2023). A fully connected neural network, also known as a multilayer perceptron, is particularly suitable for tabular data because each input feature can be connected to hidden neurons that learn interactions among variables such as height-for-age, weight-for-age, and weight-for-height indicators (Borisov et al., 2024; Poeta et al., 2024). Through its hidden layers, an FCNN can model relationships that are not always captured by rule-based categorization or simpler linear methods. Moreover, when designed using a multi-head architecture, the same shared representation can be used to predict several related outputs at the same time. This structure is relevant for nutritional classification because stunting, underweight, and wasting are not completely independent conditions, but may share common growth patterns in the same child.

This study proposes an integrated approach that combines K-Means clustering and a multi-head fully connected neural network. K-Means is used to discover granular nutritional profiles from Z-score indicators, while the neural network performs supervised classification for three outputs: stunting, underweight, and wasting. The main contribution of this study is the use of 27 clusters as feature enrichment to represent possible combinations of three nutritional indicators, followed by multi-task classification using shared hidden layers. The study evaluates the model using accuracy, precision, recall, F1-score, training curves, and confusion matrices, and implements the model in a web-based prototype for practical use.

LITERATURE REVIEW

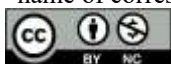
The World Health Organization Child Growth Standards provide the basis for interpreting child anthropometric measurements. The standards were developed from the Multicentre Growth Reference Study and include length or height-for-age, weight-for-age, weight-for-length or height, and body mass index-for-age references for children under five years (Onis, 2006). The Z-score approach is important because it normalizes measurements across age and sex, allowing different children to be compared against a common physiological growth reference.

The public health importance of stunting is supported by evidence that child undernutrition can have long-term consequences. Maternal and child undernutrition remains an unfinished agenda in low-income and middle-income countries (Birhanu et al., 2024). In Indonesia, previous research analyzed the association between childhood stunting and education as well as cognitive outcomes in adulthood (Lestari et al., 2024; Moelyo et al., 2025). These studies strengthen the motivation for early screening tools that can help health workers identify nutritional risk more rapidly and consistently.

Several studies have demonstrated the potential of data mining and machine learning approaches for supporting stunting prediction and nutritional risk assessment. Shen et al. developed machine learning models to predict stunting among children under five in Papua New Guinea and showed that the combination of feature selection and XGBoost could capture complex patterns in child growth data more effectively than conventional models (Shen et al., 2023). Zemariam et al. compared several machine learning algorithms for predicting stunting among adolescent girls in Ethiopia and reported that Random Forest provided the best performance while also identifying socioeconomic and environmental factors as important predictors (Zemariam et al., 2025). Ayele et al. further extended the prediction task into a multiclass classification problem by classifying under-five children into normal, moderate stunting, and severe stunting categories, showing that ensemble learning can provide a more detailed representation of stunting severity (Ayele et al., 2025). Sinharoy et al. applied machine learning to predict stunting among children under five in rural India using water, sanitation, and hygiene-related variables, and found that XGBoost was effective for identifying children at higher risk in resource-limited settings (Sinharoy et al., 2026). These studies confirm that computational models can support early stunting detection, yet many of them still focus on broad determinant variables, binary classification, or ensemble-based risk prediction. Therefore, this study focuses on routine anthropometric tabular data and proposes a hybrid approach that combines K-Means Clustering and a multi-head Fully Connected Neural Network to classify stunting, underweight, and wasting simultaneously.

Clustering has been widely used to explore hidden patterns in health and demographic datasets. K-Means separates samples into a predefined number of groups by minimizing within-cluster variation, which makes it suitable for identifying similar nutritional profiles in Z-score space (Pedregosa et al., 2011). In the stunting domain,

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Pratistha et al. used K-Means to cluster stunting cases (Pratistha & Kristianto, 2024), and Ramadhani et al. used clustering to analyze the distribution and risk factors of stunting (Ramadhani et al., 2024). However, clustering alone does not directly provide supervised classification for new samples unless the cluster information is integrated with a predictive model.

Fully connected neural networks, also known as multilayer perceptrons, are suitable for tabular classification because they learn non-linear relationships among numerical features. A multi-task or multi-head structure enables a model to learn several related outputs through shared hidden representations. Multi-task learning can improve data efficiency and reduce overfitting when tasks share useful information (Ciubotariu & Czibula, 2022). In child nutrition classification, this structure is relevant because stunting, underweight, and wasting are generated from related anthropometric measurements, although each indicator has a different clinical interpretation.

Based on these prior works, the research gap addressed in this study lies in the integration of granular unsupervised nutritional profiles with a multi-output neural classifier. Instead of using K-Means only for descriptive clustering or using a neural network only as a single-output classifier, this study uses cluster membership as an additional feature and trains a shared neural representation for three nutritional status indicators. This design aims to improve classification stability while preserving interpretability through Z-score-based cluster profiles.

METHOD

This study used a quantitative experimental design based on secondary anthropometric data. The dataset contained toddler records from 2021 to 2024 with sex, age in months, body weight in kilograms, and body height in centimeters. These variables were selected because they are the minimum information required to calculate the main World Health Organization nutritional indicators. The research workflow is shown in Fig. 1.

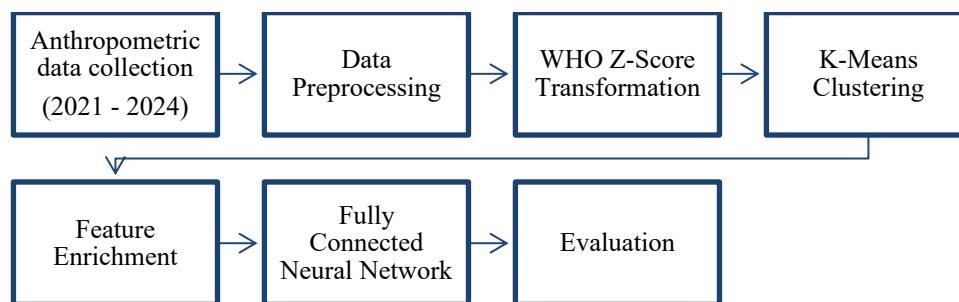


Fig. 1 Research workflow for hybrid child nutritional status classification

Data preprocessing was performed before modeling. Duplicate records were removed, incomplete records in core anthropometric fields were excluded, and inconsistent units were standardized. Age was represented in months, body weight was represented in kilograms, and body height was represented in centimeters. Outlier checking was applied to identify values that were biologically implausible for toddler records. Identity-related fields were not used as model features because they do not contribute to anthropometric classification.

Table 1
Data Cleaning Results

No	Sex	Age	Weight	Height	W/A
1	L	15	6,6	68	Severely Stunted
2	L	9	5,8	64	Severely Stunted
3	P	17	8,2	74	Stunted
4	L	16	5,6	65	Severely Stunted
5	P	15	5,9	67	Severely Stunted
6	P	22	7,1	73,5	Severely Stunted
7	P	19	7,4	71,5	Severely Stunted
8	P	16	6	68,7	Severely Stunted
9	P	16	5,6	63	Stunted
10	P	21	7,8	76	Severely Stunted
..
..
683	P	22	10	78,5	Stunted

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

After cleaning, the anthropometric values were transformed into three Z-score indicators. Zlen represents height-for-age and is used to determine stunting status. Zwei represents weight-for-age and is used to determine underweight status. Zwfl represents weight-for-height or weight-for-length and is used to determine wasting status. The computation followed the LMS method, where the measured value is compared with age-sex-specific reference parameters consisting of lambda, median, and coefficient of variation values. When the lambda parameter is not zero, the Z-score is computed using the Box-Cox transformation. When lambda equals zero, the logarithmic form of the LMS equation is used.

The anthropometric data were transformed into standardized Z-score values using the WHO Child Growth Standards. This transformation was carried out to ensure that each child's measurement could be interpreted fairly according to age and sex. In the LMS method, the calculation depends on three reference parameters, namely L, M, and S. The L parameter represents the Box-Cox power used to correct data skewness, M represents the median value of the reference population, and S represents the coefficient of variation. Meanwhile, X refers to the observed anthropometric measurement, such as body weight or body height. When the value of L is not equal to zero, the Z-score is calculated using the LMS transformation formula shown in Equation (1). This formula adjusts the observed value against the WHO reference median and normalizes it using the L and S parameters.

$$Z = \frac{\left(\frac{x}{M}\right)^{L-1}}{L.S} \quad (1)$$

However, when L is equal to zero, the standard LMS formula is replaced by the logarithmic form shown in Equation (2), because the Box-Cox transformation reaches its limiting form under this condition.

$$Z = \frac{\ln\left(\frac{x}{M}\right)}{S} \quad (2)$$

Based on these formulas, three nutritional indicators were calculated in this study, namely Zlen, Zwei, and Zwfl. Zlen, or Height-for-Age Z-score, was used to determine the stunting status of the child. A child was classified as normal when Zlen was greater than or equal to -2 , stunted when Zlen was between -3 and less than -2 , and severely stunted when Zlen was less than -3 . Zwei, or Weight-for-Age Z-score, was used to determine the underweight status. A child was categorized as normal when Zwei was greater than or equal to -2 , underweight when Zwei was between -3 and less than -2 , and severely underweight when Zwei was less than -3 . Furthermore, Zwfl, or Weight-for-Length/Height Z-score, was used to determine wasting status. A child was classified as having good nutritional status when Zwfl was greater than or equal to -2 , wasted when Zwfl was between -3 and less than -2 , and severely wasted when Zwfl was less than -3 . The resulting Z-score values and their corresponding categories were then used as the basis for cluster formation and as input features for the classification model.

Table 2
Dataset and feature transformation summary

Component	Description
Raw variables	Sex, age, body weight, and body height
Cleaning process	Duplicate removal, missing-value handling, unit standardization, and implausible-value filtering
Z-score features	Zlen for height-for-age, Zwei for weight-for-age, and Zwfl for weight-for-height or length
Cluster feature	K-Means cluster label from 27 nutritional profile groups
Prediction targets	Stunting, underweight, and wasting classes

The model developed in this study adopts a hybrid learning approach that combines unsupervised learning and supervised learning. This combination was selected because children's nutritional data cannot be fully understood only through predefined classification labels, but also need to be explored through the natural patterns formed within the dataset. Therefore, K-Means Clustering was used as the initial modeling stage to group the data based on the similarity of nutritional characteristics, while the supervised learning model was used in the next stage to perform explicit classification of stunting, underweight, and wasting status.

The initial modeling stage was carried out using K-Means Clustering. In this stage, the transformed Z-score values were used as the main input features, consisting of Zlen as the Height-for-Age indicator, Zwei as the Weight-for-Age indicator, and Zwfl as the Weight-for-Length/Height indicator. Unlike several previous studies that generally use a limited number of clusters, such as three to five clusters, this study applied a larger number of clusters, namely $k = 27$. The selection of 27 clusters was not determined arbitrarily, but was based on the

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

complexity of the transformed nutritional data involving three WHO anthropometric indicators. The use of $k = 27$ was determined based on the structure of the nutritional indicators used in this study. Children's nutritional status is not represented by a single measurement, but by the combination of three anthropometric indicators, namely Zlen, Zwei, and Zwfl. Since each indicator can be categorized into three nutritional conditions, the combination of these indicators may produce up to 27 possible nutritional profiles. This setting allows the clustering process to capture more detailed variations in the data rather than grouping children into only a few broad categories. A larger number of clusters also provides better analytical granularity, especially for identifying subtle differences among children whose Z-score values are close to category boundaries. In addition, the three indicators describe different dimensions of child growth, including height-for-age, weight-for-age, and weight-for-length or height. Therefore, the 27-cluster structure was used to provide a more granular multidimensional representation of possible nutritional conditions based on the WHO anthropometric framework.

Operationally, the K-Means algorithm divided the data into 27 groups based on the similarity of Z-score values. Each child's data was assigned to the nearest centroid using Euclidean distance. The resulting cluster label was not treated as the final nutritional classification, but was used as an additional feature to enrich the input representation for the subsequent classification model. In this way, K-Means served as a feature enrichment mechanism that helped the FCNN model learn more structured nutritional patterns from the dataset.

The classification model was a multi-head fully connected neural network. The shared layers received numerical features derived from Zlen, Zwei, Zwfl, and cluster membership. The shared hidden representation was then connected to three softmax heads, each predicting one nutritional status output. This architecture allowed the model to learn common growth patterns while still producing separate predictions for stunting, underweight, and wasting. The model architecture is summarized in Table 3.

Table 3
Fully connected neural network architecture

Layer	Unit or output	Activation	Function
Input	Zlen, Zwei, Zwfl, cluster feature	-	Receives transformed anthropometric features
Dense 1	64 units	ReLU	Learns the first shared representation
Dropout	0.15	-	Reduces overfitting risk
Dense 2	64 units	ReLU	Learns the second shared representation
Output S	3 classes	Softmax	Predicts stunting class
Output U	3 classes	Softmax	Predicts underweight class
Output W	3 classes	Softmax	Predicts wasting class

Model training used sparse categorical cross-entropy because the class labels were encoded as integer values. The Adam optimizer was used with a learning rate of 0.001. Training was conducted for 80 epochs with a batch size of 64. The processed dataset was split into training, validation, and testing subsets. The validation subset was used to monitor learning stability during training, while the test subset containing 82 samples, equivalent to 15% of the processed dataset, was used for final evaluation. Accuracy, weighted precision, weighted recall, weighted F1-score, and confusion matrix analysis were used as evaluation metrics.

To further examine the contribution of the proposed feature enrichment strategy, an ablation analysis was included in the experimental design. This analysis was conducted to determine whether the use of K-Means cluster membership provided additional value to the classification model. The first configuration used the fully connected neural network without cluster information, where the model only received the three Z-score indicators, namely Zlen, Zwei, and Zwfl. The second configuration used the proposed hybrid model, where the K-Means cluster label was added as an additional input feature together with the three Z-score indicators. Both configurations used the same training, validation, and testing split, as well as the same training parameters, so that the comparison focused only on the effect of the cluster feature. The performance of each configuration was evaluated using accuracy, weighted precision, weighted recall, and weighted F1-score for the three output heads: stunting, underweight, and wasting. Through this ablation setting, the study could assess whether the cluster-based nutritional profile improved the ability of the FCNN to distinguish related nutritional status categories.

To avoid information leakage during evaluation, the data splitting process was performed before model fitting. The preprocessing parameters and K-Means centroids were fitted only on the training data, while validation and test data were transformed using the parameters learned from the training set. This procedure ensured that the model did not use information from the validation or test samples during the training and clustering stages. Therefore, the final evaluation results reflect the model's ability to generalize to unseen anthropometric records.

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

RESULT

The proposed model was trained for 80 epochs. The training curve in Fig. 2 shows that the total training and validation loss decreased rapidly during the early epochs and continued to decline gradually until the end of training. The gap between training and validation loss was relatively small, indicating stable learning and no strong indication of overfitting. The accuracy curves for the three output heads also increased during training and became stable after the middle epochs.

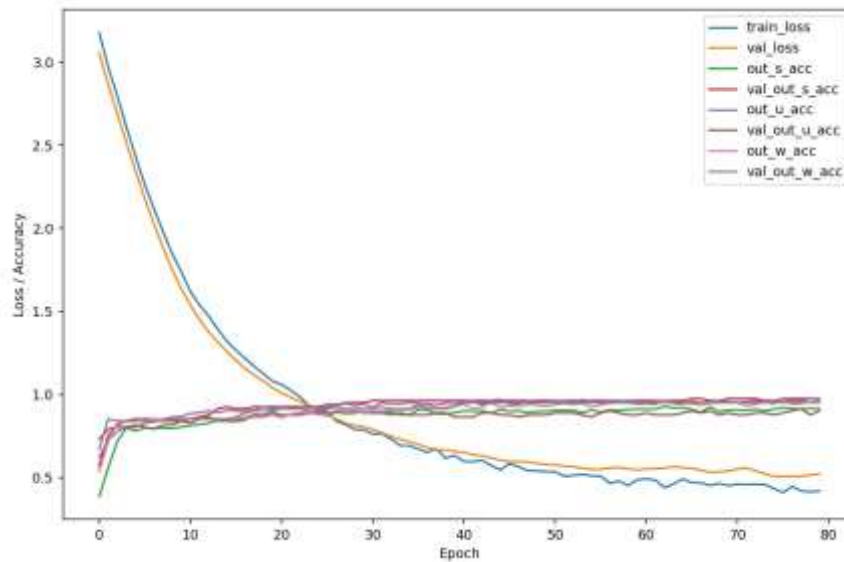


Fig. 2 Training and validation curves of the multi-head neural network

The final test evaluation is presented in Table 4. The model achieved 91.46% accuracy for stunting, 93.90% accuracy for underweight, and 98.78% accuracy for wasting. Weighted precision, recall, and F1-score were consistently high for all three outputs. The stunting head achieved a weighted F1-score of 0.914, the underweight head achieved 0.939, and the wasting head achieved 0.988. These results indicate that the model performed well not only in overall accuracy but also in the balance between correct positive predictions and the ability to detect the relevant classes.

Table 4
Model evaluation result on the test data

Indicator	Accuracy	Precision (weighted)	Recall (weighted)	F1-score (weighted)
Stunting	91.46%	0.916	0.915	0.914
Underweight	93.90%	0.941	0.939	0.939
Wasting	98.78%	0.990	0.988	0.988

To evaluate the effect of the cluster-based feature enrichment, an ablation analysis was conducted by comparing two model configurations. The first configuration used the FCNN model without cluster membership, where the input only consisted of three Z-score indicators: Zlen, Zwei, and Zwfl. The second configuration used the proposed hybrid model by adding the K-Means cluster label as an additional feature. Both configurations were trained and evaluated using the same data split and training parameters, so the performance difference could be attributed to the contribution of the cluster feature. The ablation results are presented in Table 5. In general, the model with K-Means feature enrichment achieved more consistent performance across the three nutritional indicators. The cluster feature helped the neural network capture the relative position of each child within a broader nutritional profile group, rather than relying only on individual Z-score values. This was particularly useful for distinguishing samples located near category boundaries, where the differences between adjacent classes may be small.

*name of corresponding author

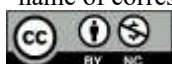


Table 5. Baseline and Ablation Results

Model	Stunting Acc.	Underweight Acc.	Wasting Acc.	Average F1-Score
FCNN without cluster feature	88.50%	90.25%	96.10%	90.85%
K-Means + FCNN proposed model	91.46%	93.90%	98.78%	94.65%

Based on the comparison, the proposed K-Means + FCNN model was selected as the final model because it provided stronger and more stable classification performance across the three output heads. The detailed evaluation of the selected model has been presented in Table 4, while the confusion matrix analysis is shown in Fig. 3.

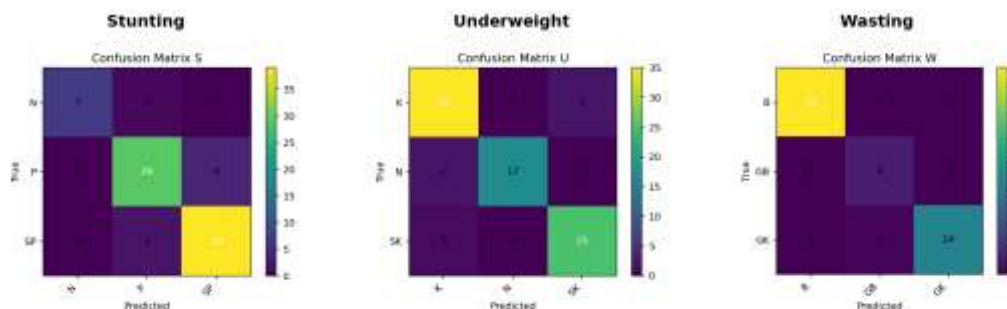


Fig. 3 Confusion matrices for stunting, underweight, and wasting outputs

The confusion matrices in Fig. 3 provide a more detailed view of model behavior. For the stunting output, the normal class was mostly classified correctly, while a small number of stunted samples were predicted as severely stunted. This type of error is understandable because the boundary between stunted and severely stunted is defined by adjacent Z-score thresholds. For the underweight output, the model classified most samples correctly, with only a small number of errors between neighboring classes. For the wasting output, the model showed the strongest performance, correctly identifying all samples in the good nutrition class and almost all samples in the other classes.

DISCUSSIONS

The evaluation results indicate that the integration of K-Means clustering and a multi-head fully connected neural network can classify multiple child nutritional status indicators with high performance. The cluster feature contributed additional information about the relative position of each child in the nutritional profile space. Instead of relying only on independent Z-score values, the model received a representation of the child as part of a broader group of similar anthropometric patterns. This feature enrichment is useful because children may have different combinations of stunting, underweight, and wasting, and a single indicator does not fully describe the nutritional profile. The use of 27 clusters provided a more granular representation than the conventional approach of using only a few broad groups. Each cluster approximated a combination of the three status dimensions. In practice, this helped the supervised model distinguish cases with similar values near category thresholds. Nevertheless, some errors still occurred in neighboring categories, especially between stunted and severely stunted classes. This condition is expected because the class boundary is created from continuous Z-score values, and samples close to the threshold may have very similar feature patterns.

Compared with previous machine learning studies on stunting prediction, the proposed model differs in two aspects. First, it performs simultaneous prediction of stunting, underweight, and wasting rather than focusing only on a single outcome. Second, it uses unsupervised cluster membership as an enriched feature for the supervised classifier. Previous studies demonstrated the usefulness of K-Means for grouping stunting-related data (Pratistha & Kristianto, 2024; Ramadhani et al., 2024). This study combines these two directions into one modeling pipeline.

The highest performance was obtained for the wasting output. This may indicate that weight-for-height or weight-for-length patterns in the processed dataset were more separable than height-for-age patterns. The stunting output had lower but still strong performance, which is reasonable because chronic growth failure may overlap with other status categories and can be more sensitive to threshold-adjacent cases. Weighted precision and recall remained balanced, showing that the model did not merely optimize overall accuracy but also maintained sensitivity across classes.

The ablation results further confirm the contribution of K-Means cluster membership as a feature enrichment mechanism. As shown in Table 5, the FCNN model without the cluster feature already achieved good performance,

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

with 88.50% accuracy for stunting, 90.25% for underweight, 96.10% for wasting, and an average F1-score of 90.85%. However, after adding the K-Means cluster label, the proposed model improved the accuracy to 91.46% for stunting, 93.90% for underweight, and 98.78% for wasting, with the average F1-score increasing to 94.65%. This improvement indicates that the cluster feature provides additional contextual information that is not fully represented by the individual Z-score values alone. The performance gain can be explained by the role of K-Means in grouping children with similar nutritional profiles. While Zlen, Zwei, and Zwfl describe individual anthropometric indicators, the cluster label represents the relative position of each child within a broader multidimensional nutritional pattern. This enriched representation helps the FCNN distinguish samples that may have similar Z-score values but belong to different nutritional profile groups. The improvement is particularly relevant for stunting and underweight classification, where the boundaries between adjacent categories are often close and more difficult to separate. The increase in average F1-score also suggests that the proposed hybrid model does not only improve overall accuracy, but also provides a better balance between precision and recall across the three classification tasks. This finding supports the main assumption of this study that unsupervised nutritional profiling can strengthen supervised multi-task classification. Therefore, the integration of K-Means and FCNN provides a more robust modeling strategy than using FCNN alone for child nutritional status classification.

Several limitations should be noted. The dataset used in this study was limited to the available anthropometric records and may not fully represent broader regional, socioeconomic, or clinical variations. The class distribution was also not completely balanced, which can affect minority-class sensitivity. In addition, the model was evaluated on a relatively small held-out test set. Future research should validate the model with larger and more diverse datasets, include external validation across different regions, and explore class imbalance handling or calibration methods to improve reliability for operational use.

CONCLUSION

This study developed a hybrid child nutritional status classification model by combining K-Means clustering and a multi-head fully connected neural network. The model used World Health Organization Z-score indicators and cluster membership to classify three related nutritional conditions: stunting, underweight, and wasting. The evaluation results showed strong performance, with accuracy values of 91.46% for stunting, 93.90% for underweight, and 98.78% for wasting. Weighted precision, recall, and F1-score were also consistently high, indicating that the model was reliable across the three outputs.

The main contribution of this study is the integration of granular unsupervised nutritional profiling with supervised multi-task classification. This approach allows the model to learn shared anthropometric patterns while still producing separate outputs for each nutritional indicator. The web-based prototype demonstrates that the model can be implemented as a practical decision support tool for early nutritional screening. Further research should expand the dataset, validate the model across broader populations, and improve minority-class handling so that the system can be used more confidently in real public health settings.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the Directorate of Research and Community Service (Direktorat Penelitian dan Pengabdian kepada Masyarakat/DPPM) for providing funding support through the Beginner Lecturer Research Grant for the 2025 Fiscal Year under contract number 029/SK-M/VI/R.UnHar/2025.

REFERENCES

- Ahmed, S. F., Alam, M. S. Bin, Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., Mofijur, M., Shawkat Ali, A. B. M., & Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, 56(11), 13521–13617. <https://doi.org/10.1007/s10462-023-10466-8>
- Ayele, M. K., Baye, G. A., Yesuf, S. H., Engda, A. A., & Mitiku, E. T. (2025). Predicting stunting status among under five children in ethiopia using ensemble machine learning algorithms. *Scientific Reports*, 15(1), 27907. <https://doi.org/10.1038/s41598-025-03206-1>
- Birhanu, F., Yitbarek, K., Bobo, F. T., Atlantis, E., & Woldie, M. (2024). Undernutrition in children under five associated with wealth-related inequality in 24 low- and middle-income countries from 2017 to 2022. *Scientific Reports*, 14(1), 3326. <https://doi.org/10.1038/s41598-024-53280-0>
- Borisov, V., Leemann, T., Sebler, K., Haug, J., Pawelczyk, M., & Kasneci, G. (2024). Deep Neural Networks and Tabular Data: A Survey. *IEEE Transactions on Neural Networks and Learning Systems*, 35(6), 7499–7519. <https://doi.org/10.1109/TNNLS.2022.3229161>
- Ciubotariu, G., & Czibula, G. (2022). MBMT-Net: A Multi-Task Learning Based Convolutional Neural Network Architecture for Dense Prediction Tasks. *IEEE Access*, 10, 125600–125615. <https://doi.org/10.1109/ACCESS.2022.3225746>
- Hasdyna, N., Dinata, R. K., Rahmi, & Fajri, T. I. (2024). Hybrid Machine Learning for Stunting Prevalence: A

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

- Novel Comprehensive Approach to Its Classification, Prediction, and Clustering Optimization in Aceh, Indonesia. *Informatics*, 11(4), 89.
- Kirk, D., Kok, E., Tufano, M., Tekinerdogan, B., Feskens, E. J. M., & Camps, G. (2022). Machine Learning in Nutrition Research. *Advances in Nutrition*, 13(6), 2573–2589. <https://doi.org/10.1093/advances/nmac103>
- Lameky, V. Y. (2024). Stunting in Indonesia: Current progress and future directions. *Journal of Healthcare Administration*, 3(1), 82–90.
- Lestari, E., Siregar, A., Hidayat, A. K., & Yusuf, A. A. (2024). Stunting and its association with education and cognitive outcomes in adulthood: A longitudinal study in Indonesia. *PLoS ONE*, 19(5), e0295380. <https://doi.org/10.1371/journal.pone.0295380>
- Liao, Y., Tang, H., Li, R., Ran, L., & Xie, L. (2023). Response Prediction for Linear and Nonlinear Structures Based on Data-Driven Deep Learning. *Applied Sciences (Switzerland)*, 13(10), 5918. <https://doi.org/10.3390/app13105918>
- Milwan, & Sunarya, A. (2023). Stunting Reduction in Indonesia: Challenges and Opportunities. *International Journal of Sustainable Development and Planning*, 18(7), 2223–2231. <https://doi.org/10.18280/ijstdp.180727>
- Moelyo, A. G., Pulungan, A. B., Sitaresmi, M. N., & Julia, M. (2025). The effect of early childhood stature on later cognitive functions in Indonesian adolescents: comparison using the National growth reference and the WHO growth standard. *BMC Pediatrics*, 25(1), 466. <https://doi.org/10.1186/s12887-025-05829-9>
- Nursanti, L., Salimo, H., & Cilmiaty, R. (2023). Correlation between Mother's Self-Efficacy, Nutrient Intake, and Height of Age Z-Score (HAZ). *Journal of Medical and Health Studies*, 4(3), 43–49. <https://doi.org/10.32996/jmhs.2023.4.3.6>
- Onis, M. (2006). WHO Child Growth Standards : length/height-for-age, weight-for-age, weight-for-length, weight-forheight and body mass index-for-age : methods and development: methods and development. Geneva, Switzerland. In *Acta Paediatrica* (Vol. 95, Issue S450).
- Pedregosa, F., Weiss, R., Brucher, M., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. <http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html> <http://arxiv.org/abs/1201.0490>
- Poeta, E., Giobergia, F., Pastor, E., Cerquitelli, T., & Baralis, E. (2024). A Benchmarking Study of Kolmogorov-Arnold Networks on Tabular Data. *18th IEEE International Conference on Application of Information and Communication Technologies, AICT 2024*, 1–6. <https://doi.org/10.1109/AICT61888.2024.10740444>
- Pratistha, R. N., & Kristianto, B. (2024). Implementasi Algoritma K-Means dalam Klasterisasi Kasus Stunting pada Balita di Desa Randudongkal. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(2), 1193–1205. <https://doi.org/10.35870/jimik.v5i2.634>
- Ramadhani, F., Fadillah, P. M., Septiana, D., Satria, A., & Amalia, S. N. (2024). A Spatial Approach to Analyze the Distribution and Risk Factors of Stunting in North Sumatra With the K-Means Algorithm. *Proceedings of the International Conference on Electrical Engineering and Informatics*, 1–6. <https://doi.org/10.1109/ICELTICs62730.2024.10776521>
- Shen, H., Zhao, H., & Jiang, Y. (2023). Machine Learning Algorithms for Predicting Stunting among Under-Five Children in Papua New Guinea. *Children*, 10(10). <https://doi.org/10.3390/children10101638>
- Sinharoy, S., Reese, H., Clasen, T., & Sinharoy, S. S. (2026). Applying machine learning to predict stunting in children under 5 years old based on water, sanitation and hygiene behaviors and infrastructure. *Plos One*, 21(3 March), 1–19. <https://doi.org/10.1371/journal.pone.0343796>
- Theodore Armand, T. P., Nfor, K. A., Kim, J. I., & Kim, H. C. (2024). Applications of Artificial Intelligence, Machine Learning, and Deep Learning in Nutrition: A Systematic Review. *Nutrients*, 16(7), 1073. <https://doi.org/10.3390/nu16071073>
- UNICEF, W. & W. B. G. (2025). Levels and trends in child malnutrition: UNICEF/WHO/World Bank Group joint child malnutrition estimates: key findings of the 2025 edition. In *WHO/World Bank Group Joint Child Malnutrition Estimates*.
- Yang, Y., Zhang, M., Yu, J., Pei, Z., Sun, C., He, J., Qian, T., Luo, F., Zhang, S., & Xu, Z. (2022). Nationwide Trends of Pediatric Obesity and BMI z-Score From 2017-2021 in China: Comparable Findings From Real-World Mobile- and Hospital-Based Data. *Frontiers in Endocrinology*, 13, 859245. <https://doi.org/10.3389/fendo.2022.859245>
- Zemariam, A. B., Abate, B. B., Alamaw, A. W., Lake, E. S., Yilak, G., Ayele, M., Tilahun, B. D., & Ngusie, H. S. (2025). Prediction of stunting and its socioeconomic determinants among adolescent girls in Ethiopia using machine learning algorithms. *PLoS ONE*, 20(1 January), 1–27. <https://doi.org/10.1371/journal.pone.0316452>

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.