International Journal Software Engineering and Computer Science (IJSECS)
5 (3), 2025, 1141-1147
Published Online December 2025 in IJSECS (http://www.journal.lembagakita.org/index.php/ijsecs)
P-ISSN: 2776-4869, E-ISSN: 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v5i3.5553.

RESEARCH ARTICLE **Open Access**

Stunting Prediction in Toddlers Using the K-Nearest Neighbor (KNN) Method Based on a Web Application at Batealit Community Health Center, Jepara

Lisa Falichatul Ibriza *
Universitas Islam Nahdlatul Ulama Jepara, Jepara Regency, Central Java Province, Indonesia.
Corresponding Email: ichaibriza66@gmail.com.

Gentur Wahyu Nyipto Wibowo
Universitas Islam Nahdlatul Ulama Jepara, Jepara Regency, Central Java Province, Indonesia.

Teguh Tamrin
Universitas Islam Nahdlatul Ulama Jepara, Jepara Regency, Central Java Province, Indonesia.

Received: September 10, 2025; Accepted: November 15, 2025; Published: December 1, 2025.

Abstract: Stunting is still a nutritional problem that exists in Indonesia and it needs immediate intervention in Jepara Regency. At the primary healthcare level, Batealit Public Health Center uses manual anthropometric recording for toddlers' growth assessment. This method can be prone to human recording errors and operational delays which hinder prompt clinical decision-making. To improve this condition, this study develops a web-based system for predicting stunting based on the K-Nearest Neighbor (KNN) algorithm. The research method was applied research with system development using the Waterfall model by processing main variables such as age, weight, and height. We tested the algorithm intensively by trying different neighbor values (k) to obtain the maximum value for accuracy, precision, and recall. From experiments, the KNN algorithm is best at k=3 with a 95.23% accuracy rate; this configuration is better compared to larger k values since they increase misclassification rates on normal and stunted categories. By porting this logic into a web interface, detection moves from being a manual task to an automated one occurring in real-time thus application becomes an essential part of decision support enabling health workers to bypass administrative delays and find stunting much faster more accurately within Batealit service area.

Keywords: Stunting; Prediction; K-Nearest Neighbor; Web Application; Anthropometry.

1. Introduction

Stunting remains a major barrier to public health improvement in Indonesia, as it is a chronic nutritional disorder that reduces the potential of human capital. More than just an issue of physical height, stunting represents a systemic failure in early childhood nutrition that undermines physical growth and cognitive development during the critical golden period of development. Empirical evidence shows that children who are stunted have much higher chances of experiencing motor and cognitive developmental disorders. These early deficits usually have long-term impacts such as reduced learning ability and delayed educational attainment, which automatically reduce economic productivity in adulthood [1]. The magnitude of this crisis can be seen from the 2022 Indonesia Nutritional Status Survey (SSGI), which recorded a national prevalence rate of 21.6%. This figure is still above the 20% tolerance threshold set by the World Health Organization (WHO), meaning that malnutrition is not an isolated case but rather a structural health emergency; hence, there is an urgent need for acceleration in early detection so that quality future human resources do not suffer irreversible risks due to stunting [2].

Although this mandate is urgent, it does not match the reality on the ground where primary healthcare facilities operate with Community Health Centers (Puskesmas). Most regions including Jepara Regency still use manual anthropometric workflows for nutritional monitoring of toddlers. Health workers take weight and height measurements by writing them down on paper charts which makes this method very prone to human error. These analog processes create risks from wrong data entry all the way to subjective misinterpretation of growth standards. It can lead to "false negatives," whereby at-risk children are wrongly classified as normal and hence do not get any intervention needed since they were never flagged as at risk in the first place. The administrative burden imposed by manual processing also creates significant operational latency that prevents immediate clinical feedback. An assessment conducted recently at Bukittinggi Community Health Center validated these inefficiencies and claimed that moving toward digital nutrition surveillance was necessary to ensure both accuracy and timeliness in monitoring processes [3].

The use of information technology in the form of machine learning architectures provides a digital answer to these analog bottlenecks. Health facilities can standardize diagnosis and eliminate subjective bias by moving from manual interpretation to algorithmic classification. Among the many predictive models available, K-Nearest Neighbor (KNN) has been identified as an especially appropriate algorithm for medical diagnostics. KNN is based on proximity — it classifies new data points based on how similar they are to existing historical records. This approach is popular because it is computationally simple and highly effective at recognizing patterns. Previous studies have shown that web-based applications using KNN for stunting detection can achieve accuracy levels above 90% consistently, validating the algorithm's performance in real-world applications [4][5].

Other algorithms like Naive Bayes or Random Forest may be powerful, but comparative studies have shown that KNN is very competitive and sometimes even more precise in certain classification tasks related to nutritional status [3]. It is non-parametric which means it does not presume any underlying distribution of the data and therefore can adapt perfectly to the specific demography of a local population. This study takes note of those technical merits by proposing the development of a web-based stunting prediction application specifically for Batealit Community Health Center in Jepara Regency. The KNN algorithm will be embedded into a user-friendly web interface with an aim to change the current manual process into one that can be diagnosed automatically and in real-time. This digital shift shall not only enhance mathematical accuracy in classifying nutritional status but also act as an important decision support system so that health workers can respond quickly and effectively against stunting.

2. Related Work

Stunting remains a chronic pediatric health problem in Indonesia with the highest consequences for the long-run human capital of the country. Stunting is a consequence of chronic malnutrition that impairs linear growth and development of the brain. Empirical studies have shown that children who are stunted are at greater risk of delayed cognitive development, reduced learning ability, and lower educational attainment in adulthood [1][2]. Tafili, Hamadi, and Spaulding (2022) further noted that it has multifactorial determinants which involve very complex interactions between nutritional intake, maternal health status, and environmental sanitation; hence long-term risks can only be mitigated by rapid and accurate early diagnosis [6].

Detection mechanisms at primary healthcare facilities such as Puskesmas for early intervention remain largely analog despite their criticality. Health workers mostly use manual anthropometric measures of weight, height, and age to determine nutritional status. Although this workflow is operationally simple, it is fundamentally flawed since it is prone to human recording errors and subjective interpretation of growth charts with great latency in data processing. Putri *et al.* (2023) contend that these operational inefficiencies often postpone an appropriate clinical decision until the condition has progressed because at-risk children will not be identified until then [3]. Hence, there must be a shift from manual recording to digital surveillance if biases and delays associated with traditional methods are to be eliminated.

The adoption of computational intelligence provides an answer to these diagnostic bottlenecks. Hasdyna, Dinata, and Fajri (2024) assert that hybrid machine learning approaches improve classification accuracy and prediction compared to manual methods significantly [7]. In the same vein, Jana, Dey, and Ghosh (2023) observed that digital health interventions have great potential for maternal and child nutrition monitoring provided primary care-level infrastructural issues are adequately addressed [8]. Information technology advancement allows big data processing toward objective data-based predictions with no human error involvement in diagnosis.

The K-Nearest Neighbor algorithm is known for its computational simplicity and high classification accuracy within machine learning algorithms. Novalina *et al.* (2025) performed a benchmark study on several algorithms and stated that KNN is very suitable for predicting stunting risk in the Indonesian population [9]. Hardt *et al*. (2023) also mentioned in their review that machine learning models are being more widely used in clinical nutrition research with good results, but integration into systems continues to be a problem [10]. KNN uses the proximity principle by finding the closest historical data to classify new data, so it is very appropriate to analyze anthropometric parameters. In addition, KNN can be implemented in a web-based architecture so that data can be accessed in real-time. This was demonstrated by Purwati and Sulistyo (2023), who created a decision support system based on KNN to help medical personnel determine and monitor stunting risk accurately [4].

3. Research Method

This study employed a quantitative research approach utilizing the K-Nearest Neighbor (KNN) machine learning algorithm to predict stunting in toddlers. The primary research data were obtained from the Batealit Community Health Center in Jepara Regency, comprising anthropometric records including age, sex, weight, and height.

1) Data Source
The dataset was collected from the official medical records of the Batealit Community Health Center, specifically covering children aged 0–59 months. Prior to analysis, the data underwent pre-processing, which involved removing duplicate entries and handling missing values to ensure the quality and integrity of the dataset.
2) Research Variables
The study defined specific variables for the classification model:
   a) Input Variables: Age (months), sex, weight (kg), and height (cm).
   b) Output Variable: Nutritional status, categorized strictly as "stunting" or "normal" in accordance with the WHO Child Growth Standards.
3) Data Analysis Method
The K-Nearest Neighbor (KNN) algorithm was applied to classify the nutritional status of toddlers. The analysis proceeded through the following stages:
   a) Data Pre-processing: Anthropometric data were normalized to ensure all variables operated on a uniform scale.
   b) Choosing *k* Values: The model was tested using several neighbor values (*e.g*., k=3, 5, 7) to identify the configuration that yielded optimal performance.
   c) Distance Calculation: The Euclidean Distance formula was utilized to quantify the similarity between data points:

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)2}$$

Where:
d(x,y) : Distance between test data \$x\$ and training data \$y\$.
x_i : The \$i\$-th attribute of the test data.
y_i : The \$i\$-th attribute of the training data.
N : Total number of attributes.

   d) Classification: The class of the test data was determined based on the majority class found among the *k* nearest neighbors.
   e) Model Evaluation: Classification accuracy was measured using the formula:

$$Accuracy = Total\ Data\ Number\ of\ Correct\ Predictions \times 100\%$$

4) System Implementation
The validated KNN model was integrated into a web-based application. The system is designed to process anthropometric input data, execute the classification using the KNN algorithm, and present prediction results in real-time to support the early detection of stunting risks.

5) Evaluation
The system's performance was evaluated through two primary mechanisms: accuracy testing of the KNN model against actual historical data, and functional testing of the web application to ensure usability for health workers. This evaluation used toddler data from the Batealit Community Health Center as a specific case study.

4. Result and Discussion

4.1 Results

The experimental phase utilized anthropometric records—specifically age, weight, height, and sex—obtained from the Batealit Community Health Center in Jepara Regency. Prior to analysis, the dataset underwent normalization to ensure all attributes operated on a uniform scale. We assessed predictive performance by testing three distinct neighbor configurations (k=3, 5, and 7). As illustrated in the graph below, the model achieved its peak performance at k=3, reaching approximately 95.2% accuracy. Increasing the neighbor count resulted in a decline in precision, with accuracy dropping to 94.4% at k=5 and 94.3% at k=7.

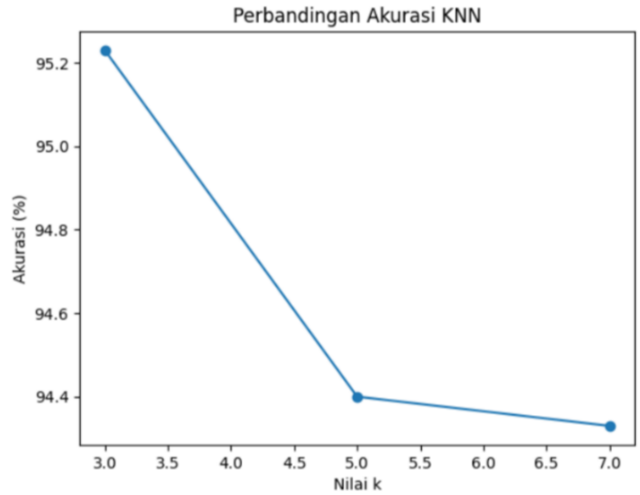

Figure 1. Comparison Graph

We further decomposed the performance of the optimal model (k=3) using a confusion matrix. The classification breakdown reveals that the system correctly identified 430 toddlers as stunted and 457 children as having a normal nutritional status. Additionally, 372 records falling outside the WHO standard age range were correctly flagged by the system.

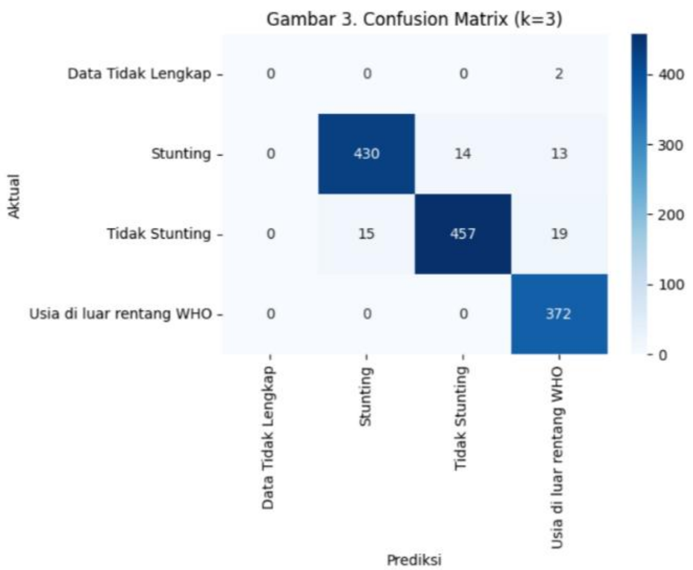

Figure 2. Confusion Matrix (K=3)

Detailed evaluation metrics are presented in the table below. The model attained an overall accuracy of 95.23%, with precision, recall, and F1-scores for the primary stunting and non-stunting categories consistently exceeding 93%. Notably, the system achieved a perfect recall rate (100%) for detecting age outliers.

=== Tabel 2. Precision, Recall, dan F1-score ===

	precision	recall	f1-score
Data Tidak Lengkap	0.000000	0.000000	0.000000
Stunting	0.966292	0.940919	0.953437
Tidak Stunting	0.970276	0.930754	0.950104
Usia di luar rentang WHO	0.916256	1.000000	0.956298
accuracy	0.952345	0.952345	0.952345
macro avg	0.713206	0.717918	0.714960
weighted avg	0.952230	0.952345	0.951562

Figure 3. Precision, Recall, and F1-Score

Following the algorithmic validation, the model was deployed within a client-server web architecture utilizing Flask for the back-end and HTML, CSS, and JavaScript for the front-end. The figure below displays the examination input interface, where administrators manually enter toddler health metrics including weight and height. Upon submission, the back-end algorithm processes these inputs to generate a nutritional status prediction.

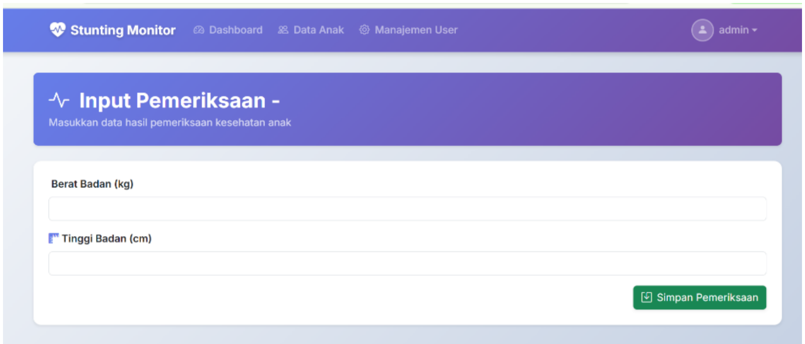

Figure 4. Toddler Data Input Page Display

The final output is visualized in the dashboard shown below. The interface employs visual indicators to differentiate status: rows highlighted in red denote toddlers detected as stunted, while rows remaining white indicate a normal nutritional status.

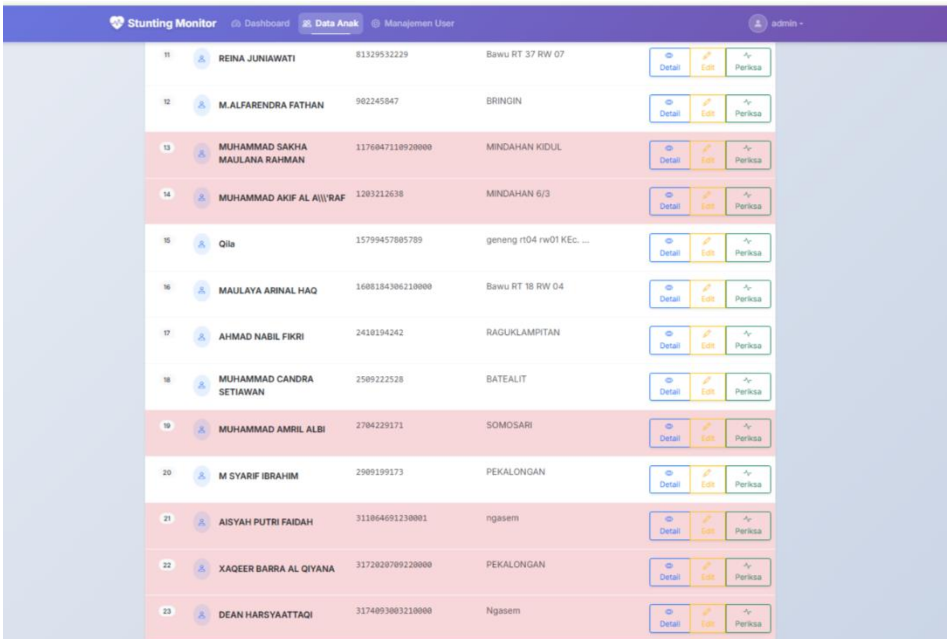

Figure 5. Nutritional Status Prediction Results Display

4.2 Discussion

The experimental outcomes indicate an inverse relationship between the number of neighbors and prediction accuracy for this specific dataset. The superior performance observed at k=3 suggests that a smaller neighborhood structure is more effective at capturing the local data patterns required for accurate stunting prediction, whereas larger neighborhoods introduce noise that dilutes precision. With an accuracy of 95.23%, the proposed KNN model demonstrates high competitiveness when benchmarked against recent studies. This result aligns closely with findings by Khansa and Gunawan (2024) and Khoirunnisa and Gunawan (2024), who reported KNN accuracies of 95.67% and 96.19%, respectively, in similar toddler datasets [12][14].

Furthermore, the stability of the KNN algorithm in this study corroborates the systematic review by Indrisari *et al*. (2025), which categorizes KNN as a stable classifier for stunting prevalence with expected accuracy ranges between 72% and 99.92% [15].Regarding comparative algorithm performance, this study confirms that KNN consistently functions as a robust classifier, particularly when compared to Naïve Bayes. Widhari *et al*. (2024) and Sinaga *et al*. (2025) both noted that KNN provides significantly better diagnostic precision compared to Naïve Bayes, making it a pragmatic choice for primary healthcare applications where computational simplicity and high accuracy must be balanced [11][16]. However, it is acknowledged that ensemble methods like Random Forest can occasionally yield marginally higher accuracy in complex datasets, as observed in studies by Putri *et al*. (2024) and Khoirunnisa and Gunawan (2024) [12][13]. Despite this, the confusion matrix in this study reveals that KNN maintains a low error rate for the stunting category, supporting the findings of Putri *et al*. (2024) who highlighted that KNN excels in Recall metrics, a critical factor in medical screening to minimize false negatives [13].From an operational perspective, the integration of these findings into a web interface directly addresses the efficiency gaps inherent in manual recording. The utilization of color-coded results, as visualized in Figure 5, serves a functional clinical purpose beyond mere aesthetics. This visual distinction enables health workers to identify high-risk children instantly, facilitating a significantly faster response protocol compared to traditional manual list reviews. Consequently, the system functions effectively as a decision support tool, combining high mathematical accuracy with practical usability to enhance stunting prevention efforts at the primary healthcare level.

5. Conclusion

This study successfully developed a web-based stunting prediction system by applying the K-Nearest Neighbor (KNN) algorithm using anthropometric data of toddlers from Batealit Community Health Center, Jepara Regency. The experimental results demonstrated that the KNN algorithm achieved a high level of accuracy, reaching 95.23% at k = 3. Evaluation using the confusion matrix, precision, recall, and F1-score also confirmed the strong performance of the model, particularly in classifying the non-stunting category and detecting children outside the WHO standard age range. The implementation of the web-based system provides several advantages, including real-time data access, ease of use for health workers, and improved speed in the early detection process. The use of color indicators in prediction results also facilitates quick identification of children requiring special attention. Therefore, the combination of the KNN method and a web-based application can serve as an effective decision support system for stunting prevention efforts at the primary healthcare level. Nevertheless, this study still has limitations, as the dataset was restricted to a single region. Future research is recommended to employ larger and more diverse datasets and to compare the performance of KNN with other machine learning algorithms to obtain more comprehensive results.

References

[1] Akbar, R. R., Kartika, W., & Khairunnisa, M. (2023). The effect of stunting on child growth and development. *Scientific Journal*, *2*(4), 153-160. https://doi.org/10.56260/sciena.v2i4.118.

[2] 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*, 1–18. https://doi.org/10.1371/journal.pone.0295380

[3] Putri, D. K., Nursal, D. G. A., Markolinda, Y., & Masrizal, M. (2023). Factors Associated with Pregnant Women's Engagement in COVID-19 Vaccination. *Jurnal Kesehatan*, *14*(3). https://doi.org/10.35730/jk.v14i3.1051.

[4] Purwati, N., & Sulistyo, G. B. (2023). Stunting early warning application using KNN machine learning method. *Jurnal Riset Informatika*, *5*(3), 373–378. https://doi.org/10.34288/jri.v5i3.550

[5] Situmeang, N., Komarsyah, E. A., Fauzi, A., & Sagala, J. (2025). Penggunaan Algoritma K-Nearest Neighbor (KNN) Untuk Mendeteksi Stunting Pada Anak. *Innovative: Journal Of Social Science Research*, *5*(3), 8401-8417. https://doi.org/10.31004/innovative.v5i3.20343.

[6] Tafili, A., Zakari, N. M., Hamadi, H. Y., & Spaulding, A. (2022). Association of hospital and community factors on the attainment of Baby-Friendly designation: A breastfeeding health promotion. *Maternal & child nutrition*, *18*(3), e13388. https://doi.org/10.1111/mcn.13388

[7] Hasdyna, N., Dinata, R. K., Rahmi, & Fajri, T. I. (2024). Hybrid Machine Learning for Stunting Prevalence: A Novel Comprehensive Approach to Its Classification, Prediction, and Clustering Optimization in Aceh, Indonesia. *Informatics, 11*(4), 89. https://doi.org/10.3390/informatics11040089.

[8] Jana, A., Dey, D., & Ghosh, R. (2023). Contribution of low birth weight to childhood undernutrition in India: evidence from the national family health survey 2019–2021. *BMC Public Health*, *23*(1), 1336. https://doi.org/10.1186/s12889-023-16160-2

[9] Novalina, N., Tarigan, I. A. A., Kameela, F. K., & Rizkinia, M. (2025). Benchmarking machine learning algorithm for stunting risk prediction in Indonesia. *Bulletin of Electrical Engineering and Informatics*, *14*(3), 2252-2263. https://doi.org/10.11591/eei.v14i3.8997

[10] Hardt, L. M., Herrmann, H. J., Reljic, D., Jaensch, P., Zerth, J., Neurath, M. F., & Zopf, Y. (2023). Are guideline recommendations on supportive nutrition and exercise therapy for cancer patients implemented in clinical routine? A national survey with real-life data. *Nutrients*, *15*(14), 3172. https://doi.org/10.3390/nu15143172.

[11] Widhari, W., Triayudi, A., & Sari, R. T. K. (2024). Implementation of Naïve Bayes and K-NN algorithms in diagnosing stunting in children. *SAGA: Journal of Technology and Information System*, *2*(1), 164–174. https://doi.org/10.58905/saga.v2i1.242

[12] Khoirunnisa, K., & Gunawan, P. H. (2024, December). Analysis of stunting prediction for toddlers in Bekasi Regency using the K-Nearest Neighbors and Random Forest algorithms. In *2024 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)* (pp. 932–936). IEEE. https://doi.org/10.1109/ICICYTA64807.2024.10913041

[13] Putri, I. P., Terttiaavini, T., & Arminarahmah, N. (2024). Analisis perbandingan algoritma machine learning untuk prediksi stunting pada anak: Comparative analysis of machine learning algorithms for predicting child stunting. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(1), 257–265. https://doi.org/10.57152/malcom.v4i1.1078

[14] Khansa, G. A. F., & Gunawan, P. H. (2024, July). Predicting stunting in toddlers using KNN and Naive Bayes methods. In *2024 International Conference on Data Science and Its Applications (ICoDSA)* (pp. 17–21). IEEE. https://doi.org/10.1109/ICoDSA62899.2024.10651676

[15] Indrisari, E., Febiansyah, H., & Adiwinoto, B. (2025). A systematic literature review on the application of machine learning for predicting stunting prevalence in Indonesia (2020–2024). *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, *14*(3), 277–283. https://doi.org/10.32736/sisfokom.v14i3.2366

[16] Sinaga, M., Fujiati, F., & Halawa, D. (2025). Designing a stunting prediction model using machine learning to support SDGs achievement in Indonesia. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, *9*(4), 2070–2079. https://doi.org/10.33395/sinkron.v9i4.15296