Green Intelligent Systems and Applications

Research Article
Volume 6(1), 2026, 133−146
https://doi.org/10.53623/gisa.v6i1.1190

Integration of Naïve Bayes-Based Stunting Status Classification and GIS Hotspot Mapping for the Identification of Priority Areas in Tomohon City, Indonesia

Eunice Emely Eurika Pitoy*, Chatreen Rindu Ceyzia Pontoh, Marike Kondoj, Herry Langi[4], Maksy Sendiang

Program Studi Teknik Informatika, Politeknik Negeri Manado, Manado, Indonesia

*Correspondence: emelypitoy04@gmail.com

SUBMITTED: 12 May 2026; REVISED: 29 May 2026; ACCEPTED: 31 May 2026

ABSTRACT: Stunting remained a public health problem that required data- and area-based monitoring so that interventions could be implemented in a targeted manner. This study aimed to develop an integrated system for classifying stunting status and identifying priority areas in Tomohon City through the combination of WHO Z-Score standards, the Naïve Bayes algorithm, prevalence calculation, and hotspot mapping based on a Geographic Information System (GIS). This study employed a Research and Development (R&D) approach consisting of needs analysis, design, implementation, testing, and evaluation stages. Toddler data were obtained from the Tomohon City Health Office, including age, sex, height or body length, weight, residential area, urban village, district, and community health center. The system was developed using MySQL, Python, PHP Framework CodeIgniter 3, and GIS. The results showed that the system was able to classify toddlers' nutritional status into normal, stunted, and severely stunted categories, calculate prevalence by urban village, and display the distribution of cases in the form of a digital map. Gaussian Naïve Bayes modeling using 970 training data points and 243 testing data points produced an accuracy of 94.7%, precision of 31.6%, recall of 33.3%, and F1-score of 32.4%. GIS hotspot visualization helped identify priority areas, although data coverage still needed to be expanded to make the results more representative.

KEYWORDS: Stunting, Naïve Bayes, WHO Z-Score, GIS, priority areas.

1. Introduction

Stunting remained one of the public health issues that required serious attention because it was directly related to children's growth quality, cognitive development, long-term productivity, and the quality of human resources. This problem was not only understood as a disorder of physical growth, but also as an indicator of chronic nutritional problems, sanitation issues, access to health services, socioeconomic conditions, and regional disparities. In Indonesia, the issue of stunting still showed variation across regions, so its management could not rely solely on general prevalence recapitulation. A data-based approach was needed to identify the individual stunting status of children under five while also showing areas with a high concentration of cases, so that interventions could be directed more accurately and effectively [1, 2].

Stunting monitoring at the regional level often still relied on manual recording or aggregate reporting, which had not fully been able to provide a rapid overview of children's nutritional status and the distribution of cases by area. In fact, anthropometric data such as age, sex, body weight, and height could be processed into more systematic classification information to support early detection. Previous studies had shown that data mining-based classification methods could be used to help identify stunting status more quickly and in a more structured manner. Naïve Bayes was one of the widely used algorithms because it had a simple probabilistic mechanism, could work with both categorical and numerical data, and was relatively easy to implement in health information systems [3, 4].

The use of Naïve Bayes in the context of stunting had been carried out by several researchers. For example, Titimeidara and Hadikurniawati applied the Naïve Bayes Classifier to classify stunting nutritional status among toddlers and showed that this method could support the data collection and classification process based on toddler attributes [3]. Another study also showed that Naïve Bayes could be implemented in a web-based expert system to assist in detecting stunting among toddlers using variables such as age, height, and body weight [4]. These findings indicated that classification algorithms could become an important part of health decision-support systems, particularly in assisting health workers and local governments in monitoring toddlers' nutritional status more efficiently.

In addition to classification, stunting also required a spatial approach because stunting cases were not randomly distributed, but might be associated with environmental characteristics, population density, access to health services, sanitation, and geographical conditions. A study by Eryando et al. showed that stunting prevalence in Indonesia had spatial variation across districts/cities and needed to be analyzed by considering spatial autocorrelation and regional intervention priorities [1]. Kamaruddin et al. also emphasized that geographical and environmental conditions might be related to the distribution of stunting, making geospatial analysis important for understanding the distribution patterns of cases at the local level [2]. This showed that the GIS approach did not only function as a visualization medium, but also as an analytical tool for determining priority areas.

In the development of information systems, Geographic Information Systems (GIS) had been widely used to map the distribution of stunting. Fathurrahman et al. developed a web-based geographic information system for mapping stunting locations in Gereneng Timur Village to display case information spatially [5]. Gobel et al. also developed a web-based geographic information system for the distribution of stunting cases in Pohuwato Regency, which was capable of displaying geographic information on stunting cases from the highest to the lowest areas and presenting prevalence at the village level [6]. Both studies showed that GIS could help local governments and health workers read regional patterns more easily compared to ordinary tabular reports.

However, most of these studies still tended to separate the classification approach for children's status from spatial mapping. Naïve Bayes-based studies mostly focused on detecting or classifying stunting status, while GIS-based studies mostly focused on visualizing distribution patterns or mapping areas. In fact, decision-making in stunting management required two types of information simultaneously: information on the individual status of children under five and information on the concentration of areas that required priority intervention. The integration of these two approaches was important because classification results could be used as input for hotspot mapping, while hotspot maps could help identify areas with higher spatial risk [7, 8].

The GIS hotspot approach was highly suitable and relevant because it could clearly show areas with high concentrations of cases. A study by Asparian et al. used spatial autocorrelation analysis to identify hotspots of stunting cases in Kerinci Regency and emphasized the importance of more specific interventions at the regional level [8]. Research by Puspitasari et al. also showed that geospatial analysis could be used to identify areas with stunting hotspot clusters among children under five, so that mapping results could serve as a basis for determining high-risk areas [9]. Therefore, GIS hotspots could strengthen classification results by providing spatial context for the classified child data.

The need to integrate machine learning and GIS was also in line with current developments, in which machine learning models could be used to predict or classify stunting based on various child and household factors. A study conducted by Chilyabanyama et al. compared several machine learning algorithms, including Naïve Bayes, to classify stunting among children under five and showed that machine learning approaches could support data-driven prediction of stunting risk [10]. Meanwhile, spatial studies indicated that stunting had regional heterogeneity, meaning that intervention strategies could not be generalized across all locations [11, 12]. These findings reinforced the need to combine classification and spatial approaches so that the analysis results became more comprehensive.

Based on the above explanation, this study was considered feasible because it contributed an approach that combined individual-level analysis and regional-level analysis. Naïve Bayes was used to classify the stunting status of children under five based on anthropometric attributes and supporting characteristics, while GIS hotspots were used to map areas with high concentrations of cases. This integration was expected to produce more applicable information for local governments, community health centers, and stakeholders in determining priority areas for stunting intervention in Tomohon City. The novelty of this study lay in combining the results of stunting status classification with spatial hotspot analysis, so that the research output was not only the status of children under five, but also a region-based priority intervention map.

2. Materials and Methods

This study employed a Research and Development (R&D) approach with structured stages consisting of needs analysis, system design, implementation, testing, and evaluation. This approach was selected because the study not only analyzed stunting data but also developed an integrated system for stunting status classification, regional prevalence calculation, Naïve Bayes modeling, and hotspot mapping based on a Geographic Information System (GIS). These stages are consistent with information system development processes that require a systematic workflow so that each function can be validated according to user requirements [13]. The developed system integrates the WHO Z-Score standard, MySQL database, Naïve Bayes algorithm, and spatial visualization to identify priority areas for stunting intervention in Tomohon City. The research data consisted of under-five children’s data obtained from the Tomohon City Health Office, which were used as experimental input for system integration.

The dataset included children's identity information, residential location, age, sex, height or body length, body weight, and growth measurement history. Administrative location data, such as village, district, and public health center coverage area, were used to link classification results with spatial information. All data were stored in a MySQL database to ensure systematic management, including data storage, retrieval, updating, processing, and presentation.

2.1. Research flow and data preprocessing.

The research workflow is presented in Figure 1, which illustrates the stages from data collection to system evaluation. The initial stage involved data collection and preprocessing. The collected data were checked for completeness, cleaned from input errors, duplicate records, and incomplete entries, and then standardized according to system requirements. After preprocessing, the system calculated the height-for-age or length-for-age Z-score based on anthropometric growth indicators. These indicators are widely used to assess linear growth in children because they reflect age- and sex-specific growth patterns and are effective for identifying chronic growth problems related to stunting [14]. Based on the computed Z-scores, toddlers were classified into three categories: normal, stunted, and severely stunted. Children were categorized as normal when the Z-score was ≥ -2 SD, stunted when -3 SD $\leq$ Z-score < -2 SD, and severely stunted when the Z-score was < -3 SD.

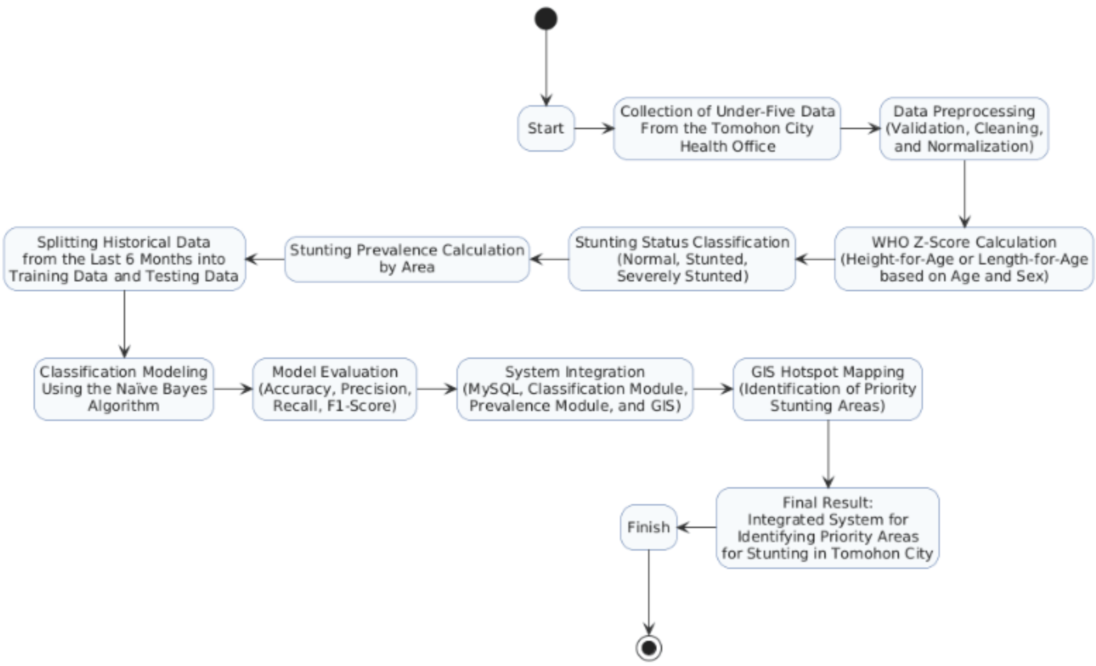


Figure 1. Research Flow

2.2. Stunting prevalence and naïve bayes classification.

After classification, the system calculated stunting prevalence for each administrative area by comparing the number of stunted and severely stunted children with the total number of examined children in that area. The formula used was:

$$\text{Stunting Prevalence} = \frac{\text{Number of Stunted and Severely Stunted Children}}{\text{Number of Measured Children}} \times 100\%$$

The resulting prevalence values were used to identify the severity of stunting across areas and served as a basis for determining priority intervention zones. Thus, the system not only generated individual nutritional status classification but also produced aggregated spatial indicators to support regional monitoring in Tomohon City. In addition to WHO Z-score-based classification, the Naïve Bayes algorithm was applied using historical data of under-five children from the previous six months. The dataset was divided into training and testing sets to build and evaluate the model. Naïve Bayes was selected because it is a supervised learning method based on Bayes' Theorem that estimates class probabilities using input attributes. Previous studies have shown that Naïve Bayes can effectively support early stunting identification by processing child-related variables and producing reliable classification outputs [15]. The class labels in this study were derived from WHO Z-score classifications, while predictor variables included age, sex, height or body length, body weight, residential area, and growth measurement history. The Z-score itself was not used as an input feature to avoid redundancy and label leakage, ensuring that the model learned patterns from raw attributes rather than pre-classified outcomes.

2.3. System implementation and GIS visualization.

The system was implemented by integrating the WHO Z-score classification module, prevalence computation, Naïve Bayes algorithm, MySQL database, and GIS-based mapping. Spatial visualization was developed using a web-based GIS approach, where the administrative map of Tomohon City was displayed interactively. The use of Leaflet-based web GIS supports interactive visualization features such as map layers, markers, and pop-up information, enabling clearer interpretation of spatial data [16]. Each administrative area was displayed based on stunting prevalence or case counts, allowing high-burden areas to be easily identified as priority intervention zones. This spatial representation improves interpretability compared to tabular data and supports decision-making based on geographic distribution patterns. Previous studies have emphasized that stunting cases often exhibit spatial clustering rather than uniform distribution, making GIS-based analysis essential [1, 8].

2.4. System evaluation.

System evaluation was conducted through functional testing and model performance assessment. Functional testing used black-box methods to verify that all system modules operated correctly, including data management, Z-score classification, prevalence calculation, database integration, and GIS visualization. The Naïve Bayes classification model was evaluated using accuracy, precision, recall, and F1-score. Accuracy measured overall correctness, while precision and recall assessed the model's ability to correctly identify stunting cases. The F1-score was used as a balanced metric because it combines precision and recall, making it suitable for evaluating classification performance, particularly in datasets with potential class imbalance [17].

3. Results and Discussion

3.1. Data collection and preprocessing.

This study used toddler data from the Tomohon City Health Office as the primary dataset for developing a stunting status classification system and GIS hotspot mapping. The data included toddler identity information, age, sex, height or body length, body weight, residential area, urban village, sub-district, and community health center working area. These variables served as the main input for stunting status classification, regional prevalence calculation, Naïve Bayes modeling, and GIS-based spatial visualization. Table 1 presents the toddler data attributes used in this study, including demographic, anthropometric, and spatial variables required for both classification and mapping processes.

Table 1. Toddler data attributes used in the study.

No	Data Attribute	Description
1	Toddler Name/ID	Toddler identity in the system
2	Gender	Male or female
3	Age	Toddler age in months
4	Height/Body Length	Main anthropometric data
5	Body Weight	Supporting measurement data
6	Urban Village	Toddler's residential location
7	Sub-district	Administrative area
8	Community Health Center	Health service working area
9	Stunting Status	System classification result
10	GIS Coordinates/Area	Basis for mapping case distribution

After preprocessing, the cleaned data were stored in a MySQL database. Using a relational database structure allowed systematic management of data, including input, updating, searching, classification, and visualization in tables, dashboards, and GIS maps. This integrated structure also enabled the linkage between individual toddler records and administrative area data, supporting spatial analysis and hotspot identification. Figure 2 illustrates the system dashboard, which integrates data management, classification output, and spatial visualization in a single interface. The dashboard allows users to access both individual stunting status results and regional distribution patterns.

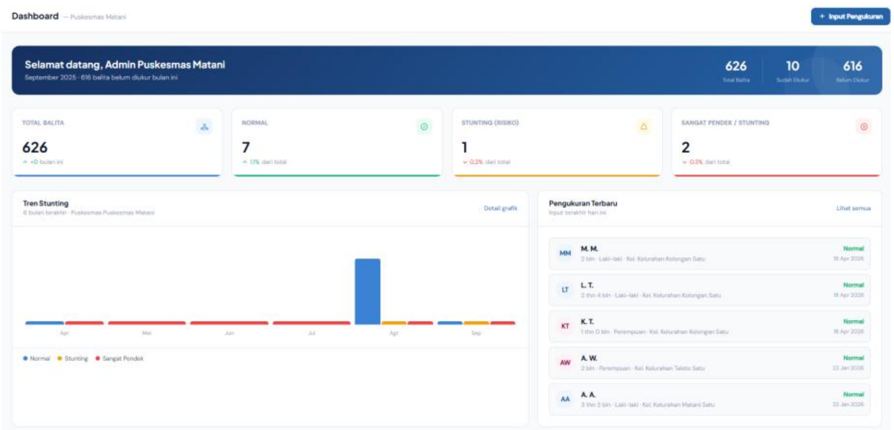


Figure 2. System dashboard.

3.2. Classification of Stunting Status Based on WHO Z-Score

The classification of stunting status was conducted using the WHO Z-score standard based on height-for-age or length-for-age indicators. The system calculated Z-score values using each child's age, sex, and height or body length data. Based on the results, children were classified into three categories: normal (Z-score ≥ −2 SD), stunted (−3 SD ≤ Z-score < −2 SD), and severely stunted (Z-score < −3 SD). This classification approach is consistent with WHO growth standards and has been widely applied in national nutrition surveillance systems [18]. Figure 3 shows the data input interface used for stunting classification. The system processes anthropometric variables entered by users and automatically computes the Z-score to determine the nutritional status category.

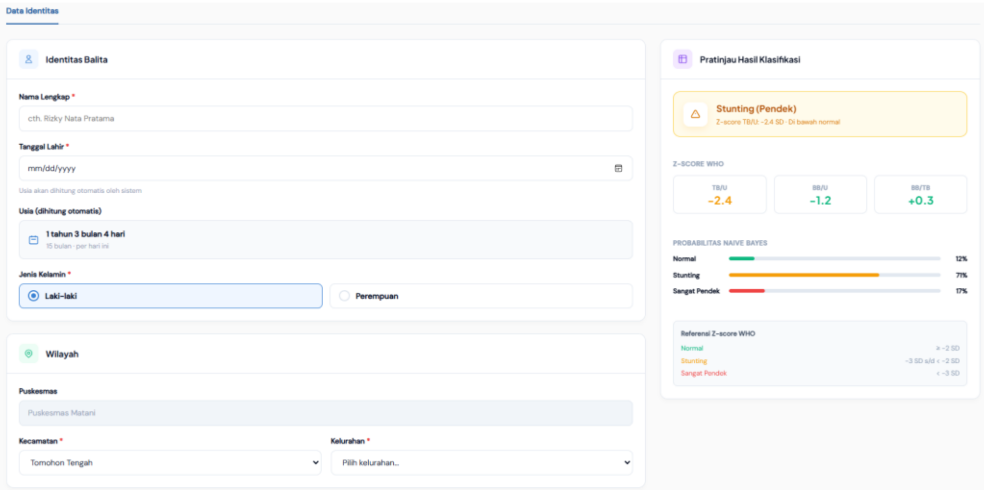


Figure 3. Data input for stunting status classification using Z-score.

Based on the system implementation, each child's data stored in the database can be processed automatically to generate classification results. This automation improves the efficiency of identifying at-risk children, particularly those categorized as stunted and severely stunted. The resulting classification outputs are then used as input for calculating the number of cases in each geographic area. Figure 4 presents the classification results generated by the system. It shows the categorized stunting status of toddlers, which serves as the basis for further prevalence calculation and spatial mapping analysis.

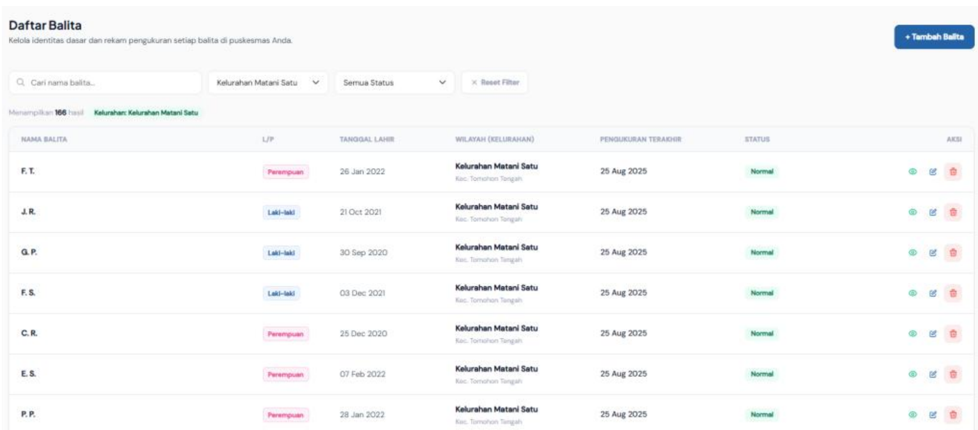
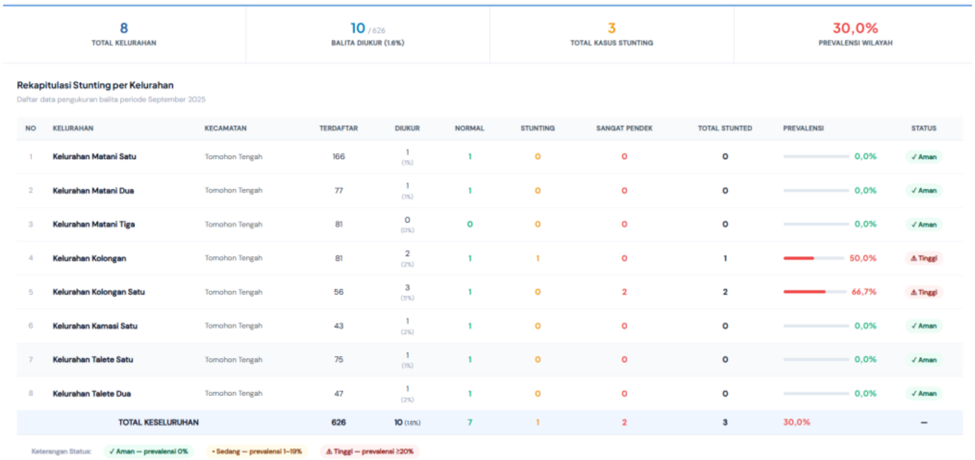


Figure 4. Classification results display.

3.3. Calculation of stunting prevalence by region.

This study used toddler data from the Tomohon City Health Office as the primary dataset for developing a stunting status classification system and GIS hotspot mapping. The data included toddler identity information, age, sex, height or body length, body weight, residential area, urban village, sub-district, and community health center working area. These variables served as the main basis for stunting status classification, regional prevalence calculation, Naïve Bayes modeling, and GIS-based spatial visualization. The detailed attributes used in this study are summarized in Table 1, which presents the structured dataset used for system development and analysis. After preprocessing, all data were stored in a MySQL database to ensure structured management, including data input, updating, retrieval, classification, and visualization. The integrated database also enabled linkage between individual toddler records and administrative area information for spatial analysis. The overall system workflow and dashboard interface used to manage and visualize the data are shown in Figure 5, which illustrates the integration of data management, classification, and GIS-based visualization modules.


8 TOTAL KELURAHAN | 10 / 626 BALITA DIUKUR (1.6%) | 3 TOTAL KASUS STUNTING | 30,0% PREVALENSI WILAYAH

Rekapitulasi Stunting per Kelurahan

Daftar data pengukuran balita periode September 2025

NO	KELURAHAN	KECAMATAN	TERDAFTAR	DIUKUR	NORMAL	STUNTING	SANGAT PENDEK	TOTAL STUNTED	PREVALENSI	STATUS
1	Kelurahan Matani Satu	Tomohon Tengah	166	1 (1%)	1	0	0	0	0,0%	✓ Aman
2	Kelurahan Matani Dua	Tomohon Tengah	77	1 (1%)	1	0	0	0	0,0%	✓ Aman
3	Kelurahan Matani Tiga	Tomohon Tengah	81	0 (0%)	0	0	0	0	0,0%	✓ Aman
4	Kelurahan Kolongan	Tomohon Tengah	81	2 (2%)	1	1	0	1	50,0%	⚠ Tinggi
5	Kelurahan Kolongan Satu	Tomohon Tengah	56	3 (5%)	1	0	2	2	66,7%	⚠ Tinggi
6	Kelurahan Kamasi Satu	Tomohon Tengah	43	1 (2%)	1	0	0	0	0,0%	✓ Aman
7	Kelurahan Talete Satu	Tomohon Tengah	75	1 (1%)	1	0	0	0	0,0%	✓ Aman
8	Kelurahan Talete Dua	Tomohon Tengah	47	1 (2%)	1	0	0	0	0,0%	✓ Aman
	TOTAL KESELURUHAN		626	10 (1.6%)	7	1	2	3	30,0%	–

Keterangan Status: ✓ Aman – prevalensi 0% • Sedang – prevalensi 1–19% ⚠ Tinggi – prevalensi ≥20%

Figure 5. Display of prevalence calculation results.

Based on the experimental data from the stunting recapitulation by urban village for the September 2025 period, the system recorded 8 urban villages in Tomohon Tengah District with a total of 626 registered children under five. Of this number, 10 children had been measured, or approximately 1.6%. The classification results showed that 7 children were classified as normal, 1 child as stunted, and 2 children as severely stunted. Therefore, the total number of stunted cases was 3 children, resulting in an overall regional prevalence of 30.0%. Most urban villages were categorized as safe, with a prevalence of 0.0%, namely Matani Satu, Matani Dua, Matani Tiga, Kamasi Satu, Talete Satu, and Talete Dua. Meanwhile, two urban villages were classified as high-risk areas, namely Kolongan with a prevalence of 50.0% and Kolongan Satu with a prevalence of 66.7%. These two areas serve as examples of priority areas because they have a higher proportion of stunted children compared to the other urban villages. The identification of high-prevalence and hotspot areas is important because spatial analysis can support the acceleration of stunting prevention by directing interventions toward locations with greater regional risk [1].

3.4. Classification Modeling Using Naïve Bayes.

The classification modeling of stunting status was carried out using the Python programming language with the Gaussian Naïve Bayes algorithm. Toddler measurement data were stored in a MySQL database and were then processed for model training and testing. This system was integrated with a web-based application developed using the PHP Framework CodeIgniter 3 (CI3), allowing the classification results and model evaluation to be displayed directly through the system interface. The dataset used consisted of 970 training records and 243 testing records, with the main features including the child's age, sex, body weight, and body height. The model evaluation results are presented in Figure 6, which shows the system interface displaying performance metrics and model outputs. The results indicate that the model achieved an accuracy of 94.7%, precision of 31.6%, recall of 33.3%, and F1-score of 32.4%. The system also generated model performance visualizations, including an F1-score graph, to help users interpret the evaluation results. Although the accuracy was relatively high, the F1-score indicated that the model performed better in recognizing the normal category than the stunted and severely stunted categories. This condition was consistent with previous research showing that class imbalance in toddler nutritional status classification can affect recall and F1-score, particularly for minority classes; therefore, model evaluation should not rely solely on accuracy [19]. Consequently, the Naïve Bayes results were used as decision support, while the final stunting status determination still referred to the WHO Z-Score standard.

.

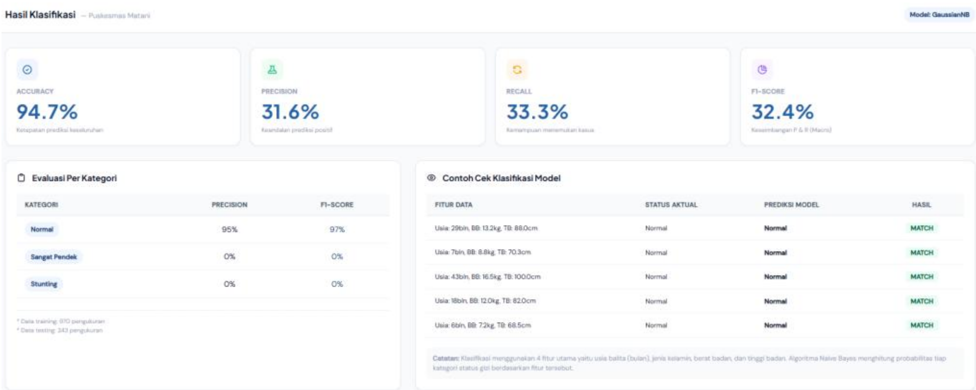


Figure 6. Display of Model Evaluation Results

3.5. GIS hotspot mapping results.

The classification results and stunting prevalence calculations were visualized using GIS-based hotspot mapping on a digital map. In the map display, each urban village was marked with a colored indicator representing the spatial distribution of stunting cases. The information displayed in each marker included the name of the urban village, sub-district, number of normal children under five, number of stunted children, number of severely stunted children, and the prevalence value. For example, in Kolongan Satu Urban Village, the system showed 1 normal child under five, 0 stunted children, and 2 severely stunted children, with a prevalence value of 3.6%. The GIS hotspot mapping output is presented in Figure 7, which illustrated the spatial distribution of stunting cases across urban villages in Tomohon City.

.

Figure 7. GIS Hotspot Mapping. Green markers indicate low-risk areas, yellow markers indicate moderate-risk areas, red markers indicate high-risk areas, blue markers represent health center or administrative reference points, and pop-up information displays urban village name, sub-district, number of normal children, number of stunted children, number of severely stunted children, and prevalence percentage.

This GIS visualization made it easier to observe the spatial distribution of stunting cases compared to tabular data alone. Through the map, areas with stunted and severely stunted cases could be directly identified based on their geographic location. Therefore, GIS hotspot mapping functioned as a supporting tool for identifying priority areas in stunting intervention programs in Tomohon City. Recent geospatial studies have shown that GIS can effectively identify hotspot areas of undernutrition among children under five, including stunting, thereby supporting location-based public health interventions [20]. The integration of classification results, prevalence data, and spatial information made the system more informative for community health centers and local governments in determining area-based intervention priorities.

3.5. Discussion.

The results of the study showed that the integration of WHO Z-Score classification, Naïve Bayes modeling, prevalence calculation, and GIS hotspot mapping was able to produce a more structured stunting monitoring system. WHO Z-Score-based classification was used as the main basis for determining children's status into normal, stunted, and severely stunted categories. This classification was relevant because stunting assessment commonly refers to anthropometric indicators of length-for-age or height-for-age, where children with a length/height-for-age value below -2 SD were categorized as stunted, while those below -3 SD were categorized as severely stunted [21]. This approach was also supported by recent evidence showing that child stunting is commonly measured using height-for-age values below two standard deviations from the WHO Child Growth Standards median [22].

The prevalence calculation results showed that the system could compare stunting conditions across regions. In the trial data, Kolongan Subdistrict and Kolongan Satu Subdistrict were categorized as high-prevalence areas, so they could be used as examples of priority intervention locations. This finding showed that stunting analysis was not sufficient when conducted only at the individual level, but also needed to be linked to regional conditions because stunting cases may form specific spatial patterns [1]. A recent geospatial study in

Tulang Bawang Regency also found that stunting incidence among toddlers showed spatial variation across areas, indicating that GIS-based analysis can support the identification of high-risk locations and area-based intervention priorities [9]. Spatial analysis among children under five in Ethiopia also showed that stunting, wasting, and underweight were not evenly distributed across regions, reinforcing the importance of area-based analysis for identifying priority intervention locations [23].

The Gaussian Naïve Bayes modeling produced an accuracy of 94.7%, while the F1-score remained low at 32.4%. This condition indicated that the model performed better in recognizing the normal category than the stunted and severely stunted categories. This may have been influenced by class imbalance in the dataset; therefore, Naïve Bayes in this study was more appropriately used as decision support rather than as the sole basis for stunting classification [10]. Similar findings were reported by Sinaga et al., who found that imbalanced child growth data could bias classification models toward the majority normal class and reduce their sensitivity in recognizing minority categories such as stunted and severely stunted children [24].

GIS hotspot mapping helped present classification and prevalence results in the form of a digital map. This visualization made it easier for community health centers and local governments to identify areas with higher case concentrations, allowing interventions to be directed more accurately. The use of GIS in stunting mapping was also in line with previous research showing that web-based geographic information systems can present the distribution of stunting cases and prevalence in a more informative way [6]. Similar spatial studies emphasized that stunting prevalence may form geographical clusters, making hotspot mapping useful for identifying priority areas and supporting more targeted public health planning [25]. A spatial autocorrelation study in Kerinci Regency also found that stunting hotspots appeared in specific areas and shifted across observation years, indicating the need for continuous spatial monitoring and village-level intervention planning [8].

4. Conclusions

This study produced an integrated system to support the identification of priority stunting areas in Tomohon City through a combination of WHO Z-Score-based classification, Naïve Bayes modeling, regional prevalence calculation, and GIS hotspot mapping. The developed system was capable of processing toddler data, determining stunting status categories, calculating prevalence at the urban village level, and displaying case distribution in the form of a digital map. The integration of Python, MySQL, PHP Framework CodeIgniter 3, and GIS made data processing more structured and accessible through the system interface. The results of the study showed that the system could assist community health centers and local governments in assessing stunting conditions more quickly, both at the individual and regional levels. Naïve Bayes modeling provided additional information to support decision-making, while the primary classification still referred to the WHO Z-Score standard. GIS hotspot visualization also helped identify areas with higher prevalence, allowing it to be used as an initial basis for determining stunting intervention priorities. However, the findings were still influenced by limited measurement coverage, so the system requires more complete and representative data for stronger generalization. For future research, it is recommended to use larger and more evenly distributed datasets across all urban villages to obtain more representative results. In addition, the Naïve Bayes model could be compared with other algorithms such as Decision

Tree, Random Forest, Support Vector Machine, or K-Nearest Neighbor. System development could also be directed toward real-time data integration, monthly trend analysis, and region-based intervention recommendations.

Author Contributions

Eunice Emely Eurika Pitoy contributed to the conceptualization, system development, data processing, analysis, and manuscript preparation. Chatreen Rindu Ceyzia Pontoh contributed to data collection, methodology development, and manuscript review. Marike Kondoj contributed to validation, interpretation of results, and manuscript revision. Herry Langi contributed to supervision, technical evaluation, and critical review of the manuscript. Maksy Sendang contributed to research supervision, final manuscript review, and approval of the submitted version. All authors have read and approved the final manuscript.

Data Availability

The data used in this study were obtained from the Tomohon City Health Office and are not publicly available due to institutional data access restrictions and privacy considerations. Data may be made available from the corresponding author upon reasonable request and with permission from the relevant institution.

References

[1] Eryando, T.; et al. (2022). Spatial Analysis of Stunting Determinants in 514 Indonesian Districts/Cities: Implications for Intervention and Setting of Priority. *Geospatial Health*, *17*(1), 1–7. https://doi.org/10.4081/gh.2022.1055.

[2] Jusni, J.; Arfiani, A.; Kamaruddin, M. (2023). Geospatial Analysis on Stunting Prevalence and Strategies. *Comprehensive Health Care*, *7*(1), 18–28. https://doi.org/10.37362/jch.v7i1.963.

[3] 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. *International Journal of Environmental Research and Public Health*, *22*, 449. https://doi.org/10.3390/ijerph22030449.

[4] Kusumaningrum, R.; Indihatmoko, T.A.; Juwita, S.R.; Hanifah, A.F.; Khadijah, K.; Surarso, B. (2020). Benchmarking of Multi-Class Algorithms for Classifying Documents Related to Stunting. *Applied Sciences*, *10*, 8621. https://doi.org/10.3390/app10238621.

[5] Ahmed, K.Y.; Agho, K.E.; Page, A.; Arora, A.; Ogbo, F.A.; on behalf of the Global Maternal and Child Health Research Collaboration (GloMACH). (2021). Mapping Geographical Differences and Examining the Determinants of Childhood Stunting in Ethiopia: A Bayesian Geostatistical Analysis. *Nutrients*, *13*, 2104. https://doi.org/10.3390/nu13062104.

[6] Gobel, S.; Abdillah, T.; Padiku, I.R. (2023). Sistem Informasi Geografis Sebaran Penderita Stunting Berbasis Web di Kabupaten Pohuwato. *Diffusion Journal of Systems Information Technology*, *3*(1), 11–21. https://doi.org/10.37031/diffusion.v3i2.19775.

[7] Lonang, S.; Normawati, D. (2022). Klasifikasi Status Stunting Pada Balita Menggunakan K-Nearest Neighbor Dengan Feature Selection Backward Elimination. *Jurnal Media Informatika Budidarma*, *6*(1), 49. https://doi.org/10.30865/mib.v6i1.3312.

[8] Asparian, A.; Wisudariani, E.; Syukri, M.; Putri, C.I. (2024). Spatial Autocorrelation Analysis to Identify Hotspots of Stunting Cases in Kerinci Regency. *Jurnal Bidan Cerdas*, *6*(1), 1–10. https://doi.org/10.33860/jbc.v6i1.3480.

[9] Puspitasari, Y.; Srw, D.W.; Darmawan, A.; Tugiyono, T. (2025). Geospatial Analysis of Stunting Incidents In Toddlers In Tulang Bawang District Lampung Province. *TPM—Testing, Psychometrics, Methodology in Applied Psychology*, *32*(4), 1475–1484. Available online: https://tpmap.org/submission/index.php/tpm/article/view/3663 (accessed on unknown).
[10] Chilyabanyama, O.N.; et al. (2022). Performance of Machine Learning Classifiers in Classifying Stunting Among Under-Five Children in Zambia. *Children*, *9*(7), 1082. https://doi.org/10.3390/children9071082.
[11] Girma, B.; Sasahu, L.D.; Rahman, A. (2025). Spatial distribution of stunting among breastfeeding children in Sub-Sahara Africa. *PLOS ONE*, *20*(6), e0325812. https://doi.org/10.1371/journal.pone.0325812.
[12] Tamir, T.T.; Tekeba, B.; Mekonen, E. G.; Zegeye, A.F.; Gebrehana, D.A. (2024). Spatial heterogeneity and predictors of stunting among under five children in Mozambique: a geographically weighted regression. *Frontiers in Public Health*, *12*, 1–17. https://doi.org/10.3389/fpubh.2024.1502018.
[13] Al Fajar, M.; Dar, M.H.; Rohani, R. (2022). Application of Waterfall Model in Development of Family Planning Participants Information System. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, *6*(2), 679–686. https://doi.org/10.33395/sinkron.v7i2.11387.
[14] Kartasurya, M.I.; et al. (2023). Determinants of length for age Z scores among children aged 6–23 months in Central Java, Indonesia: a path analysis. *Frontiers in Nutrition*, *10*, 1031835. https://doi.org/10.3389/fnut.2023.1031835.
[15] Melyani, S.; Harahap, S.Z.; Irmayanti, I. (2024). Prediction of Stunting in Toddlers Combining the Naive Bayes Method and the C4.5 Algorithm. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, *8*(2), 1160–1168. https://doi.org/10.33395/sinkron.v8i2.13651.
[16] Saputra, M.R.W.A.; Wismarini, T.D. (2025). WebGIS Fasilitas Kesehatan di Kabupaten Demak dengan Leaflet dan Web Mapping Interaktif Client-Side. *Jurnal Teknologi Informasi dan Multimedia*, *7*(3), 587–600. https://doi.org/10.35746/jtim.v7i3.751.
[17] Sujon, K.M.; Hassan, R.; Choi, K.; Samad, M.A. (2025). Accuracy, precision, recall, f1-score, or MCC? empirical evidence from advanced statistics, ML, and XAI for evaluating business predictive models. *Journal of Big Data*, *12*(1), 268. https://doi.org/10.1186/s40537-025-01313-4.
[18] Arief, Y.S.; Yunita, F.C.; Efendi, F.; Murti, F.A.K.; Pradipta, R.O.; McKenna, L. (2025). Social and Environmental Determinants of Childhood Stunting in Indonesia: National Cross-Sectional Study. *JMIR Pediatrics and Parenting*, *8*, e68918. https://doi.org/10.2196/68918.
[19] Kamil, M.I.; Wibowo, A.P. (2024). Implementation of the Naive Bayes Classifier Algorithm for Classifying Toddler Nutritional Status. *Journal of Applied Informatics and Computing*, *8*(2), 421–427. https://doi.org/10.30871/jaic.v8i2.8669.
[20] Charkos, T.G.; Arero, G.; Abdo, M. (2025). Geospatial distribution and determinants of undernutrition among children under five in Oromia Regional State, Ethiopia. *Journal of Nutrition and Metabolism*, *2025*(1), 5556781. https://doi.org/10.1155/jnme/5556781.
[21] Octavius, G.S.; Pardede, C.S. br; Thandy, C.C.; Fisca, C.A.L.; Juliansen, A. (2022). Comparison of Indonesian growth reference chart and world health organization child growth standard in detecting stunting: A systematic review and meta-analysis of 15,874 children. *Journal of Clinical Research in Pediatric Endocrinology*, *14*(1), 96. https://doi.org/10.4274/jcrpe.galenos.2021.2021-8-12.
[22] Karlsson, O.; Kim, R.; Moloney, G.M.; Hasman, A.; Subramanian, S.V. (2023). Patterns in child stunting by age: A cross-sectional study of 94 low-and middle-income countries. *Maternal & Child Nutrition*, *19*(4), e13537. https://doi.org/10.1111/mcn.13537.
[23] Kuse, K.A.; Debeko, D.D. (2023). Spatial distribution and determinants of stunting, wasting and underweight in children under-five in Ethiopia. *BMC Public Health*, *23*(1), 641. https://doi.org/10.1186/s12889-023-15488-z.

[24] 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.

[25] Sipahutar, T.; Eryando, T.; Budhiharsana, M.P. (2022). Spatial analysis of seven islands in Indonesia to determine stunting hotspots. *Kesmas: Jurnal Kesehatan Masyarakat Nasional*, *17*(3), 228. https://doi.org/10.21109/kesmas.v17i3.6201.