ComTech: Computer, Mathematics and Engineering Applications, 17(1), June 2026, 11−22 P-ISSN: 2087-1244
DOI: 10.21512/comtech.v17i1.13732 E-ISSN: 2476-907X

Explainable Machine Learning Models SHAP-based for Feature Importance Affecting Stunting Prevalence

Asysta Amalia Pasaribu[1]*; Nur Fitriyani Sahamony[2]; Khairil Anwar Notodiputro[3]; Bagus Sartono[4]

[1,2,3,4]Study Program of Statistics and Data Science, IPB University, Bogor, Indonesia 16680
[1]Statistics Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480
[2]Digital Business Study Program, Faculty of Business and Social Science, Binawan University, Jakarta, Indonesia 13630
[1]asysta.amalia@binus.ac.id; [2]kawaiisahamony@apps.ipb.ac.id; [3]khairil@apps.ipb.ac.id; [4]bagusco@apps.ipb.ac.id

Received: 10th June 2025/ **Revised:** 27th October 2025/ **Accepted:** 24th November 2025

How to Cite: Pasaribu, A. A., Sahamony, N. F., Notodiputro, K. A., & Sartono, B. (2026). Explainable machine learning models SHAP-based for feature importance affecting stunting prevalence. *ComTech: Computer, Mathematics and Engineering Applications, 17*(1), 11−22. https://doi.org/10.21512/comtech.v17i1.13732

***Abstract* -** Stunting is a form of chronic nutritional deficiency in toddlers and remains a major public health concern due to its impact on child growth and development. Efforts to reduce its prevalence continue to be strengthened in Indonesia, particularly in Sumatra Province. This study aims to evaluate the accuracy of a logistic regression model and three machine learning models—decision tree, random forest, and support vector machine (SVM)—in classifying stunting prevalence. The response variable is the prevalence of stunting among toddlers and is categorized into two classes: exceeding the national target and not exceeding it, based on the 2024 national threshold. Although classification models can provide accurate predictions, they often lack interpretability. Therefore, this study applies the SHAP method to the best-performing machine learning model to identify the key factors influencing stunting. The use of Shapley values is justified through the uniqueness theorem, which establishes it as the only attribution method that satisfies desirable fairness properties. SHAP values explain the model by referencing both the trained model and the underlying data. The results show that the random forest model achieves the highest accuracy (90.00%) and outperforms the other models. SHAP analysis reveals that Underweight is the most influential predictor contributing to stunting prevalence in Sumatra Province. These findings highlight the importance of machine learning interpretability in supporting policy decisions to reduce stunting.

Keywords: feature importance, logistic regression, explanaible machine learning, SHAPE value, stunting prevalence

I. INTRODUCTION

Stunting is a condition of growth failure in children caused by chronic malnutrition, resulting in a child's height being shorter than that of their peers. Stunting remains a serious problem worldwide (Rifada et al., 2023), particularly in poor and developing countries such as Indonesia (Ashari et al., 2023). Stunting is one of the major challenges the Indonesian government faces in developing and implementing a national prevention strategy. Stunting is a condition of reduced height in children caused by malnutrition that persists over a long period. This condition results in a child's height being shorter than that expected for their age. Efforts to reduce its prevalence in Indonesia are currently being intensified to align with global targets, particularly the national target, which aims to accelerate stunting reduction so that the prevalence among toddlers decreases to 19.4% in 2024. According to Purnamasari et al. (2022), the effects of stunting on children persist in both the near and distant future. Short-term impacts include impaired or delayed brain development, decreased intelligence quotient (IQ), and a weakened immune system, which increases susceptibility to infections and diseases. Therefore, stunting management, particularly in Indonesia, requires a comprehensive understanding of the factors that influence its prevalence.

Several studies on stunting are conducted by previous researchers using various statistical methods. Research on stunting using logistic regression models is carried out by Asmare et al. (2025), Fadmi et al. (2025), Juniarti et al. (2025), and Kassie and Asgedom (2025). These studies demonstrate the usefulness of logistic regression in identifying significant determinants of

*Corresponding Author

stunting. Researchers also apply spatial regression models to determine factors influencing stunting, as shown in studies by Asgedom et al. (2024), Fahrani et al. (2025), Falah et al. (2025), Girma et al. (2025), and Muhaimin et al. (2025).

Furthermore, because child stunting is influenced by complex interactions among various factors, traditional linear models such as logistic regression may not adequately capture potential nonlinear relationships. In this context, SHAP-based Explainable Machine Learning (EXAI) models help identify complex patterns that conventional approaches cannot detect. Methods for model interpretation, such as global feature importance or partial dependence plots, tend to provide only a general overview of feature influence on the model and do not consistently explain each feature's contribution to individual predictions or the direction of influence (positive or negative). Conventional methods such as global feature importance, PDP/ICE, and linear model coefficients are often inadequate when applied to complex machine learning models because they primarily offer global insights rather than local (individual-level) explanations and do not effectively handle feature interactions or model nonlinearity. As a result, interpreting the direction and magnitude of a specific feature's contribution remains challenging.

Therefore, the decision to use SHAP in this study arises from several methodological advantages that make it highly relevant for interpreting machine learning models in public health research. First, SHAP provides a quantitative and theoretically grounded method for attributing feature contributions to each prediction, drawing directly on the principles of the Shapley value in cooperative game theory. This approach ensures that every feature's contribution is allocated fairly, consistently, and without bias, fulfilling key axioms such as efficiency, symmetry, and additivity. Second, SHAP not only provides precise local explanations at the individual level but also enables the aggregation of these explanations to obtain a coherent global understanding of model behavior. This dual capability ensures that insights derived at the micro level remain aligned with broader model patterns and enhances interpretability across different analytical scales. Third, SHAP is inherently model-agnostic, meaning it applies across a wide range of algorithms—including decision trees, ensemble methods, and deep neural networks—without requiring modifications to model structure or training procedures. This flexibility makes SHAP particularly valuable in comparative studies such as this one, where multiple machine learning models are evaluated. Altogether, these strengths justify the use of SHAP as a robust and versatile interpretability tool that supports transparent decision-making in stunting research and other complex health-related modeling tasks.

Explainable Machine Learning becomes increasingly popular among researchers. This approach is widely applied across various fields of study. In the health sector, Explainable Machine Learning is used in studies by Houssein et al. (2025), Shifa et al. (2025), Sun et al. (2020), and Ullah et al. (2025). In psychology, Explainable Machine Learning is applied by Fedyk and Ray (2023), Henninger et al. (2023), Rehman et al. (2025), and Uban et al. (2022). The application of Explainable Machine Learning in the financial sector is reported by Arsenault et al. (2025), Mohsin and Nasim (2025), and Shah et al. (2024).

One technique in Explainable Machine Learning for identifying influential factors is SHAP. The Shapley value provides a fair method for distributing a collective gain among a group of collaborating players, and this concept underpins the SHAP framework. Just as the Shapley value allocates contributions among players, SHAP uses this principle to interpret machine learning models, which are often regarded as black boxes. Based on this explanation, the purpose of this study is to determine the accuracy of the logistic regression model and the machine learning approaches of decision trees, random forests, and support vector machines (SVMs). Furthermore, the SHAP method is applied to the best-performing machine learning model to identify the main risk factors for stunting cases in Sumatra.

II. METHODS

The most prevalent form of malnutrition is stunting (protein–energy/micronutrient deficiency), which affects toddlers' fetal growth, pregnancy nutrition, and maternal size, both before birth and in the early postnatal period. Sudiman states that stunting in toddlers serves as an indicator of long-term nutritional status and provides insight into broader socioeconomic conditions, particularly during the first two years of a child's life. This study uses secondary data, specifically raw data from the 2023 Indonesian Health Survey (SKI) report on the prevalence of stunting in Sumatra Province. In addition, data on influencing variables are obtained from the Central Bureau of Statistics (BPS). The Indonesian Health Survey (SKI) 2023 integrates the Basic Health Research (Riskesdas) and the Indonesian Toddler Nutrition Status Survey (SSGI). The SKI 2023 dataset is conducted to assess health development achievements over the past five years in Indonesia and to measure annual trends in toddler nutritional status from 2019 to 2024. The data generated provide an overview of health status at both national and district/city levels.

The unit of analysis in this study is toddlers suffering from stunting in Sumatra Province. The acceleration of stunting reduction among toddlers remains a government priority program, as outlined in the National Medium-Term Development Plan (RPJMN) for 2020–2024. The national target for 2024 is to reduce stunting prevalence to 14%. Therefore, this study applies the national stunting target threshold of 14% for 2024 as a reference in the analysis.

The dependent variable in this study is the prevalence of stunting at the district/city level in

Sumatra Province. Stunting status is categorized into two groups: cases that do not meet the national target and cases that meet the national target. Individuals diagnosed with stunting are identified based on physical examinations and supporting tests conducted by healthcare professionals. The independent variables used in this study are obtained from the Central Bureau of Statistics (BPS) and are presented in Table 1. All districts and cities in Nanggroe Aceh Darussalam, North Sumatra, Riau, South Sumatra, Bangka Belitung Islands, Bengkulu, Jambi, and Lampung provinces are included in this study.

The analysis methods used in this study include descriptive and inferential analyses, as well as machine learning techniques such as Support Vector Machine, Random Forest, and Decision Tree, along with an explainable model using SHAP (Shapley Additive Explanations). Descriptive analysis provides a general overview of stunting prevalence in Sumatra Province, Indonesia. Inferential analysis is conducted using a binary logistic regression model to examine the independent variables that influence stunting prevalence in 2023, with reference to the national target in Sumatra Province, Indonesia. The national stunting reduction target for 2024 is set at 14%, and this threshold serves as a benchmark for inferential and classification analyses.

In this study, the machine learning methods are applied both with and without SHAP. The approach without SHAP focuses solely on evaluating model performance, such as classification accuracy, without explaining the underlying reasons for the model's predictions. In contrast, the SHAP-based approach emphasizes model interpretability by identifying the contribution of each predictor variable to the prediction results. Therefore, this study compares the outcomes of machine learning models with and without SHAP to highlight the added value of explainability.

The binary logistic regression model is a data analysis technique used to determine the relationship between a binary response variable (*y*) and a set of predictor variables (*x*). In this study, the logistic regression model analyzes stunting prevalence in 2023 in Sumatra Province using eight predictor variables listed in Table 1. The mathematical form of the logistic regression model applied in this study is presented in Equation (1).

$$\frac{P(Y=1)}{1-P(Y=1)} = \beta_0 + \sum\nolimits_{i=1}^{7} \beta_i X_i \qquad (1)$$

β_0 : Intercept

$\beta_1, \beta_2, ..., \beta_7$: Regression coefficients for independent variables

$x_1, x_2, ..., {}_x\beta_7$: Independent Variable

Next, when using machine learning techniques to support decision-making, model interpretation remains crucial. The machine learning models used in this study include Random Forest, Decision Tree, and Support Vector Machine, and this study explores three different classification algorithms. Decision trees are constructed using rule-based structures. Decision trees are structured algorithms that describe features in a tree-like form and operate based on a set of rules. Each node represents a feature, each connection between nodes represents a decision rule, and the leaf nodes represent the output. Decision trees are often viewed as flowcharts, where predictions are made at the leaf nodes that mark the end of the decision process, which begins at the root node. As a result, decision trees function as decision-support tools and can also be visualized as tree-like plots that show predictions derived from a series of feature-based splits..

Furthermore, another classification method used in this study is the random forest model, which consists of a collection of decision trees and incorporates

Table 1 Variables of Dependent and Independent

Variables	**Category/Unit**
Dependent Variable	
Stunting Prevalence in accordance with National Target of 14% (Y)	1 = Not yet reached target 0 = Achieve the target
Independent Variable	
Wasting (X1)	Percentage (%)
Underweight (X2)	Percentage (%)
Poor people (X3)	Percentage (%)
Gross Regional Domestic Product per Capita (X4)	Trillion Rupiah Indonesia
Households that have adequate sanitation (X5)	Percentage (%)
Number of health workers (X6)	Person
Proportion of Ever-Married Women Aged 15-49 Who Gave Birth to a First Live Child Under 20 Years of Age by District/City (X7)	Percentage (%)

randomness to reduce the risk of overfitting. In this approach, multiple decision trees are combined, bootstrap samples are generated with replacement, and nodes are split based on the optimal division selected from a random subset of features. This ensemble mechanism improves model robustness and predictive performance. Random Forest is considered one of the most effective classification methods because it requires minimal data preprocessing while maintaining high accuracy. In addition, a Support Vector Machine (SVM) is trained using a learning algorithm that constructs an optimal decision boundary. This learning system employs a hypothesis space of linear functions in a high-dimensional feature space and applies a learning bias derived from statistical learning theory (Shah et al., 2024)

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

The selection of metrics is very important for evaluating a model. This study addresses a classification problem. The metrics commonly used are accuracy, precision, and recall, which are defined by Equations (2), (3), and (4), respectively. These metrics are defined based on the confusion matrix. The term TP (True Positive) represents the number of correct predictions when the model predicts that an observation belongs to the positive class and the prediction is indeed positive. The term TN (True Negative) represents a correct negative prediction, while FP (False Positive) refers to an incorrect prediction in which a negative observation is incorrectly classified as positive, and FN (False Negative) refers to an incorrect prediction in which a positive observation is incorrectly classified as negative.

Table 2 Performance Metrics

	Predictive Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

However, because accuracy typically favors the dominant class, this metric is limited and may not accurately reflect model performance when the data are unbalanced. This limitation arises because the values in the confusion matrix can be disproportionately influenced by the majority class. Therefore, the confusion matrix presented in Table 2 provides a more detailed explanation of model performance. The results of each machine learning algorithm, prior to the application of SHAP, are evaluated using performance metrics derived from the confusion matrix..

This section examines one approach to evaluating the collective techniques employed, specifically game theory, in the stunting case in Sumatra Province. In this section, the rationale and key features for utilizing game theory to explain prediction models are presented. These results display Shapley values and their role in enhancing model transparency and prediction accuracy. In game theory, Shapley values are also referred to as SHAP values. The two main components of these values are players and the game, where "players" represent the model's features and the "game" represents the prediction outcome produced by the model. Shapley values determine the extent to which each participant contributes to the game, and in this study, SHAP values quantify each feature's contribution to the model's output by applying this concept to the stunting case.

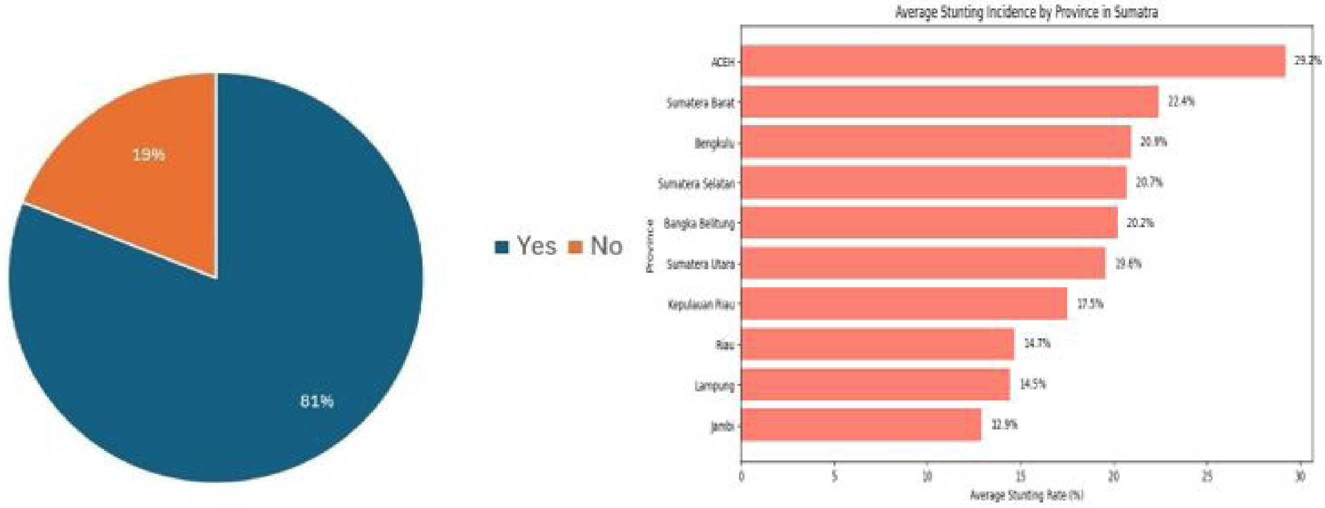
Each player's contribution is calculated by considering every possible coalition of players, or every possible combination of i features, where i ranges from 0 to n, the total number of available features. The prediction outcome is influenced by the order in which features are added to the model. This approach is introduced by Lloyd Shapley in 1953. The resulting contributions are referred to as "Shapley values," which are denoted as phi values. Then, p represents the prediction generated by the complex model; the Shapley value for a specific feature i, as expressed in Equation (5), is defined with respect to a subset S of the total set of n features.

$$\emptyset_i(p) = \sum_{S \subseteq \setminus\{i\}} \frac{|S|!\,(n - |S| - 1)!}{n!}\left[p(S \cup \{i\}) - p(S)\right] \quad (5)$$

By calculating the difference between the model's forecast and the actual value, this formula helps determine the significance or impact of a feature when feature i is excluded from the model. Equation (5) is theoretically computed using the data; however, this calculation is complex and difficult to perform directly. Therefore, Equation (5) is operationalized and visualized through the Shapley value plot. The effect of a feature is essentially reflected as the change in the model's estimated prediction when that feature is included or removed. Furthermore, SHAP values are applied whenever an input–output model is analyzed and there is a need to understand the reasoning behind the model's decisions. For example, SHAP values explain why the model predicts that stunting prevalence exceeds the national target for 2024 in Sumatra Province, Indonesia.

III. RESULTS AND DISCUSSIONS

Based on the SKI 2023 dataset of stunting cases in Sumatra Province, Figure 1 shows that 81% of provinces record stunting prevalence rates above 14%. This condition indicates that these provinces do not



(a) (b)

Figure 1 (a) Pie Chart and (b) Bar Chart of Stunting Prevalence in Sumatera Province

Table 3 Statistic Descriptive of Stunting Prevalence in Sumatera Province

Province	Min	Max	Mean	Standard deviation
Naggroe Aceh Darussalam (NAD)	15.40	40.20	29.21	5.52
North Sumatera	5.70	33.80	19.55	7.64
West Sumatera	14.70	33.70	22.43	5.64
Riau	7.60	23.00	14.68	4.82
Riau Island	15.20	21.60	17.51	2.60
South Sumatera	7.80	33.10	20.71	7.89
Kepulauan Bangka Belitung	17.30	23.20	20.21	1.93
Bengkulu	6.70	28.60	20.93	6.77
Jambi	4.10	23.70	12.94	6.23
Lampung	5.00	24.60	14.45	5.91

meet the national target, while only 19% of provinces in Sumatra meet the target. Information from the Coordinating Ministry for Human Development and Culture reports that the national prevalence of stunting in 2023 is 21.5%. Therefore, Indonesia sets a target to achieve a stunting prevalence rate of 14% in 2024. Sumatra is one of the Indonesian islands with a relatively high prevalence of stunting, particularly in Nanggroe Aceh Darussalam (NAD) Province.

Descriptive statistics for stunting cases in Sumatra Province are presented in Table 3. The descriptive measures include the minimum, maximum, mean, and standard deviation of stunting prevalence across districts and cities. Based on Figure 1 (see Appendices), the data illustrate the overall ranking of stunting prevalence in Sumatra in 2023 from the highest to the lowest. The sequence of provinces is Nanggroe Aceh Darussalam (NAD), West Sumatra, Bengkulu, South Sumatra, Bangka Belitung Islands, North Sumatra, Riau Islands, Riau, Lampung, and Jambi.

The lowest and highest prevalence rates in Nanggroe Aceh Darussalam Province are 15.40% and 40.20%, respectively. This result indicates that Nanggroe Aceh Darussalam Province does not achieve the national target for stunting prevalence. The province has 23 districts and cities with stunting prevalence levels that exceed the national target. South Aceh Regency records the highest stunting prevalence in Aceh. Aceh Tamiang Regency ranks second with a stunting prevalence of 35.9%, followed by Aceh Singkil Regency, which ranks third with a prevalence of 34.1%. The lowest prevalence of stunting among toddlers is observed in Gayo Lues Regency. Banda Aceh City ranks 21st in terms of stunting prevalence within Aceh Province.

The lowest and highest prevalence rates in North Sumatra Province are 5.70% and 33.80%, respectively. This finding indicates that several regencies and cities in North Sumatra meet the national stunting target, while others still exceed it. According to Table 4 (see Appendices), seven regencies and cities achieve the national target. Deli Serdang Regency records the highest stunting prevalence in North Sumatra, indicating the need for targeted intervention in this area.

Table 4 List of Names of Districts/Cities that Achieved Targets

Province	Number of Districts/ Cities	Number of districts/ cities that achieved the target	List of names of districts/cities that achieved the target
Naggroe Aceh Darussalam	23	0	Not available
Sumatera Utara	33	7	Asahan, Sibolga, Tebing Tinggi, Labuanbatu Utara, Pematang Siantar, Medan, Tanjung Balai
Sumatera Barat	19	0	Not available
Riau	12	5	Indragiri Hulu, Siak, Pelalawan, Pekanbaru, Kampar
Kepulauan Riau	7	0	Not available
Sumatera Selatan	17	2	Ogan Komering Ulu Timur, Lahat
Kepulauan Bangka Belitung	7	0	Not available
Bengkulu	10	1	Kota Bengkulu
Jambi	11	6	Bungo, Kota Jambi, Muaro Jambi, Batang Hari, Kerinci, Sarolangun, Kota Sungai Penuh
Lampung	15	7	Bandar Lampung, Tulang Bawang Barat, Lampung Selatan, Pesawaran, Tulang Bawang, Metro, Mesuji

The lowest prevalence of stunting among toddlers is observed in Tanjungbalai Regency. Medan City ranks eighth among districts and cities in North Sumatra for stunting cases. The lowest and highest prevalence rates in West Sumatra Province are 14.70% and 33.70%, respectively. This result indicates that West Sumatra Province does not meet the national stunting target. Mentawai Islands Regency records the highest prevalence in West Sumatra. West Pasaman Regency ranks second in West Sumatra with a stunting prevalence of 29.7%. South Solok Regency shows the lowest prevalence in West Sumatra. Padang City ranks eighth in West Sumatra in terms of stunting prevalence.

The lowest and highest prevalence rates at the regency/city level in Riau Province are 7.60% and 23.00%, respectively. Kuantan Singingi Regency records the highest stunting prevalence in Riau Province. Kepulauan Meranti Regency ranks second in Riau Province with a prevalence of 19.6%. The lowest prevalence of stunting among toddlers in Riau Province is found in Kampar Regency. Pekanbaru City ranks eleventh in stunting prevalence within Riau Province.

The lowest and highest prevalence rates in the Riau Islands Province are 17.30% and 23.20%, respectively. Based on Table 4, no district or city in the Riau Islands Province achieves the national stunting target. Bintan Regency records the highest stunting prevalence in the Riau Islands Province. Lingga Regency ranks second with a stunting prevalence of 20.5%. The lowest prevalence of toddler stunting in this province is observed in the Anambas Islands Regency. Tanjungpinang City ranks sixth in terms of stunting prevalence in the Riau Islands Province.

The lowest and highest prevalence rates in South Sumatra Province are 7.80% and 33.10%, respectively. North Musi Rawas Regency records the highest stunting prevalence in South Sumatra Province. Empat Lawang Regency ranks second in South Sumatra with a stunting prevalence of 32.6%. Ogan Komering Ilir Regency ranks third in the province with a prevalence of 32.5%. The lowest prevalence of stunting among toddlers in South Sumatra Province is found in Lahat Regency.

The lowest and highest prevalence rates at the regency/city level in Bangka Belitung Islands Province are 17.30% and 23.20%, respectively. Bangka Regency records the highest stunting prevalence in Bangka Belitung Islands Province. Belitung Regency ranks second in the province with a prevalence of 20.80%. Pangkalpinang City ranks third in Bangka Belitung Islands Province in terms of stunting prevalence. The lowest prevalence of stunting among toddlers in this province is observed in East Belitung Regency.

The lowest and highest prevalence rates at the regency/city level in Bengkulu Province are 6.70% and 28.60%, respectively. Rejang Lebong Regency records the highest prevalence of stunting in Bengkulu Province. Mukomuko Regency ranks second with a stunting prevalence of 27.10%. The lowest prevalence of stunting among toddlers in Bengkulu Province is found in Bengkulu City.

The lowest and highest prevalence rates in Jambi Province are 4.10% and 23.70%, respectively. Tanjung Jabung Timur Regency records the highest stunting prevalence in Jambi Province. Tebo Regency ranks second in the province with a prevalence of 22.70%. Jambi City ranks sixth among districts and cities in Jambi Province. The lowest stunting prevalence among toddlers in Jambi Province is observed in Sungai Penuh City.

The lowest and highest prevalence rates at the district/city level in Lampung Province are 5.00% and 24.60%, respectively. West Lampung Regency records the highest stunting prevalence in Lampung Province. These figures indicate that stunting remains

Table 5 Parameter Estimation Results of the Binary Logistic Regression Model

Variable	Coefficient	Standard Error	z-value	p-value
Constant	-8.0480	2.726	-2.952	0.003
Wasting	-0.6278	0.163	-3.841	0.000
Underweight	1.0798	0.178	6.062	0.000
Poor People	0.2816	0.109	2.582	0.010
Gross Regional Domestic Product per Capita	-0.0551	0.026	-2.117	0.034
Households that have adequate sanitation	-0.0199	0.023	-0.853	0.393
Number of health workers	0.0004	0.000	0.929	0.353
Proportion of Ever-Married Women aged 15-49 who gave birth to a first	1.1748	2.610	0.450	0.653

Table 6 Evaluation Metric Values Based on Confusion Matrix

Model	Accuracy	Sensitivity	Specificity	Precision
Logistic Regression	0.8400	0.7600	0.9200	0.9050
Decision Tree (DT)	0.8800	0.8333	0.9231	0.9090
Random Forest (RF)	0.9000	0.8400	0.9600	0.8940
Support Vector Machine (SVM)	0.8800	0.8333	0.9231	0.9091

Table 7 SHAP Values and Features Importance Per Province

No.	Province	Feature Importance	SHAP Value
1	Aceh	X_4	0.050457
2	Bangka Belitung	X_2	0.065067
3	Bengkulu	X_2	0.053173
4	Jambi	X_2	0.022113
5	Kepulauan Riau	X_2	0.050257
6	Lampung	X_1	0.017603
7	Riau	X_4	0.041926
8	Sumatera Barat	X_2	0.090757
9	Sumatera Selatan	X_4	0.046935
10	Sumatera Utara	X_2	0.018422

$$Logit(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = -8.0480 - 0.6278X_1 + 1.0798X_2 + 0.2816X_3 - 0.0551X_4 - 0.0199X_5 + 0.0004X_6 + 1.1748X_7 \quad (6)$$

a significant public health concern across several districts and cities in Lampung Province.

The results of the logistic regression model using stunting prevalence data from Sumatra Province in 2023 are presented in Table 5. Table 5 reports the parameter estimates obtained from the binary logistic regression model for stunting cases in Sumatra Province. Based on the results shown in Table 5, the estimation equation for the logistic regression model, as expressed in Equation (6), is determine as presented below.

Based on Table 5, the independent variables that significantly influence stunting prevalence in Sumatra Province in 2023 are Wasting, Underweight, Poor Population, and Gross Regional Domestic Product (GRDP). These results indicate that when Wasting and GRDP values decrease, stunting prevalence increases. Conversely, increases in the Underweight variable and the Poor Population variable are also associated with higher stunting prevalence. Table 6 presents the evaluation metric values derived from the confusion matrix. The calculated Shapley values are then summarized and presented in Table 7. Among the provinces, Nanggroe Aceh Darussalam records the highest Shapley value, indicating a strong contribution to the model's predictions.

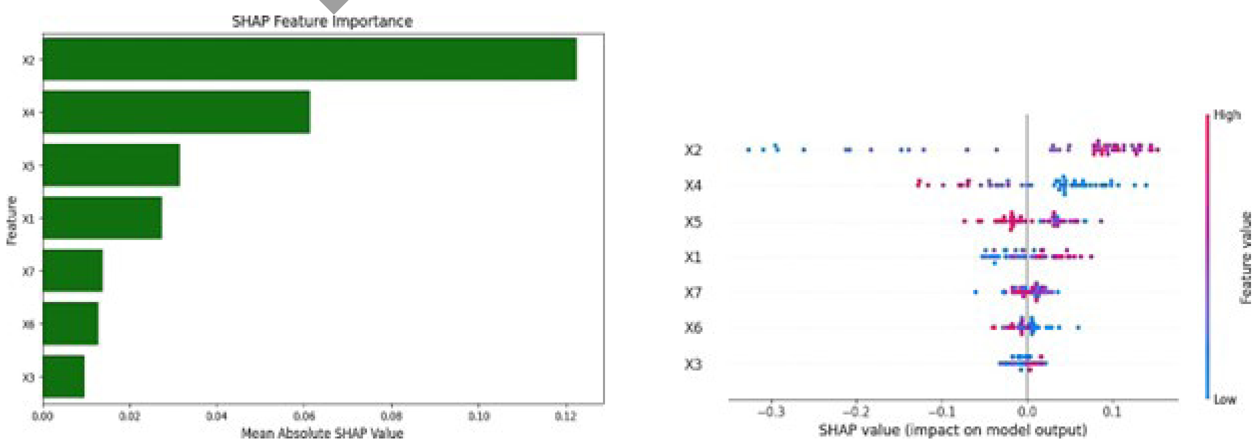
Meanwhile, other explanatory variables, such as the percentage of households with proper sanitation, the number of health workers, and the proportion of ever-married women aged 15–49 years who have given birth to their first live child, show relatively weaker effects. By examining the Random Forest model, which is identified as the best-performing model among the four models, SHAP values are calculated to interpret feature contributions. The SHAP feature importance framework is straightforward, as the importance of each feature is computed by averaging its SHAP values across observations. The most influential features are then ranked in descending order and visualized in a feature importance plot. The feature importance of the Random Forest model for stunting prevalence in Sumatra Province is displayed using a conventional bar chart, as shown in Figure 2.

Based on Figure 2, Underweight (X_2) has a greater impact than the other variables, suggesting that changes in this factor produce a more pronounced effect on stunting prevalence. When the Underweight (X_2) value increases, stunting prevalence in Sumatra Province also increases. As illustrated in Figure 2, higher Gross Regional Domestic Product per capita (X_4) is associated with a reduction in stunting prevalence. In contrast, variables such as the proportion of ever-married women aged 15–49 years who have given birth to their first child, the number of health workers, and the number of poor people exhibit relatively smaller effects on stunting prevalence. This finding indicates that variations in these variables do not substantially influence the model's predictions. As shown in Figure 2, in the SHAP summary plot, the y-axis represents the feature values, while the x-axis represents the Shapley values. The color gradient reflects the magnitude of the effect, where red indicates higher feature values and blue indicates lower values, and larger SHAP values correspond to stronger feature impacts on the prediction.

The features are ranked by predictive power, and the graphs show which features most strongly affect the model's predictions. Each feature contributes either positively or negatively to the model output. In addition, because the variables strongly correlate with the target variable, SHAP serves as an effective tool with clear advantages for variable selection. The plot in Figure 2 shows that Underweight (X_2), Gross Regional Domestic Product per capita (X_4), and Households with adequate sanitation (X_5) are the most important features, because both high and low values of these variables are strongly associated with high and low SHAP scores. In contrast, other variables, including the proportion of ever-married women aged 15–49 who have given birth to their first child, the number of health workers, and the proportion of poor people, are less influential, as their SHAP values are closer to zero and therefore exert a smaller effect on the model. Table 5 presents the parameter estimates obtained from the logistic regression analysis.

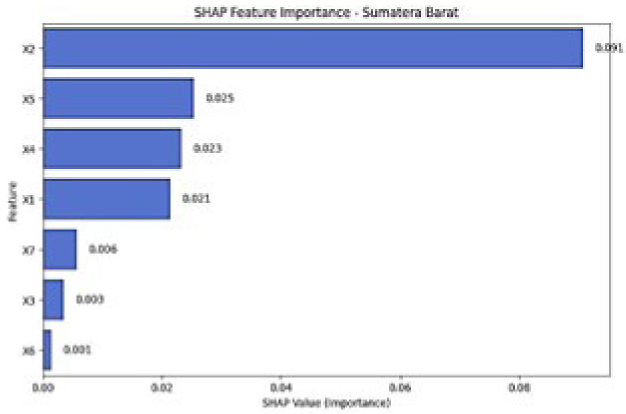
As shown in the upper rows of the SHAP summary plot, higher values of attributes such as Underweight, Gross Regional Domestic Product per capita, and Households with adequate sanitation exhibit a strong positive association with stunting prevalence. The color scale represents the magnitude of the feature values, while positive SHAP values are displayed along the x-axis. These patterns indicate that increases in these attributes are associated with higher predicted stunting prevalence. The visual representation reinforces how feature magnitude and direction jointly influence the model's predictions.

Based on the SHAP values shown in Figure 3 and Table 7, the three most important features overall are Gross Regional Domestic Product per capita (X_4), Underweight (X_2), and households with adequate sanitation (X_5). In Bangka Belitung Province, Figure 4 shows that the three most important features influencing stunting cases are Underweight (X_2), the number of health workers (X_6)), and the proportion of poor people (X_3). For Bengkulu Province, the three key features are Underweight (X_2), Gross Regional Domestic Product per Capita (X_4), and Households

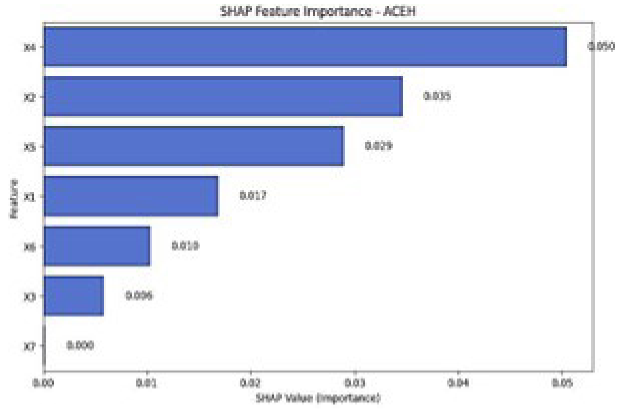


(a) (b)

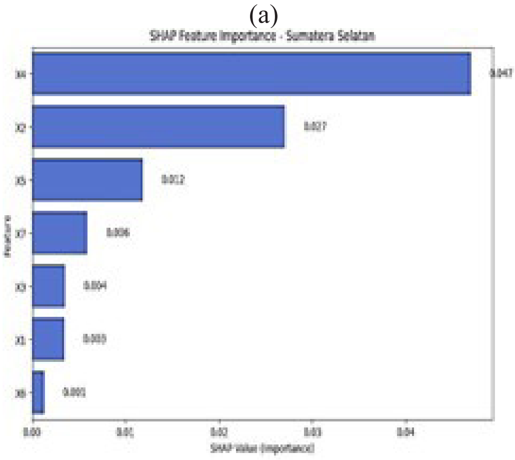
Figure 2 (a) Traditional Feature Importance SHAP Plots and (b) Shapley Value Plot



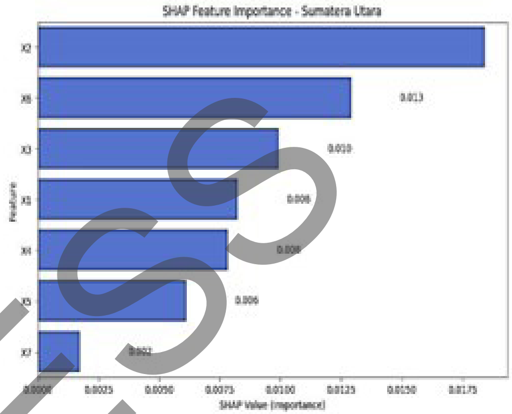
(a)



(b)

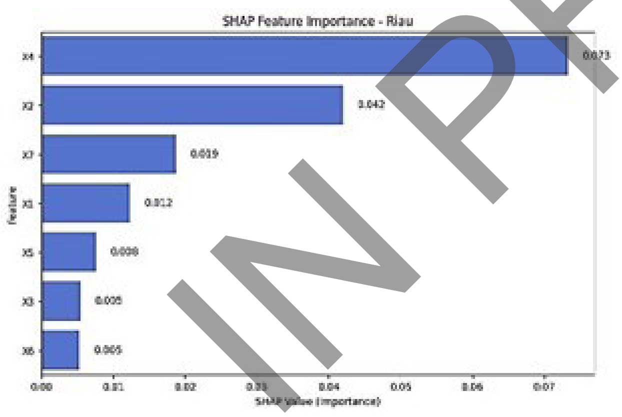


(c)

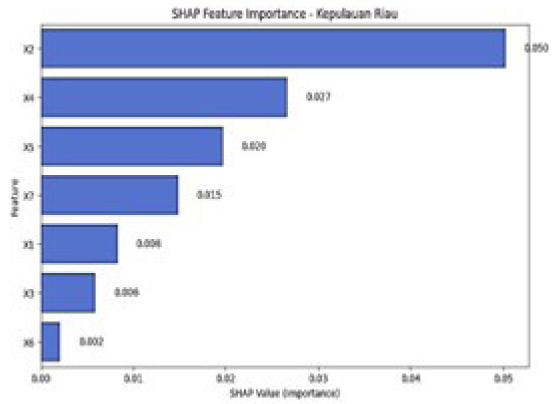


(d)

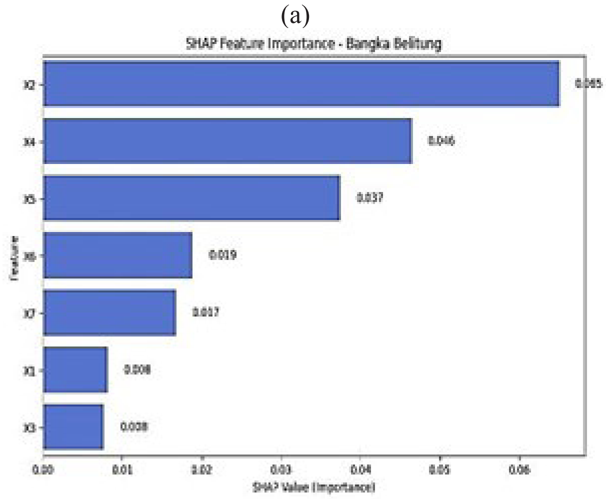
Figure 3 Traditional SHAP Plot by Province: (a) Aceh (b) West Sumatera
(c) North Sumatera (d) South Sumatera



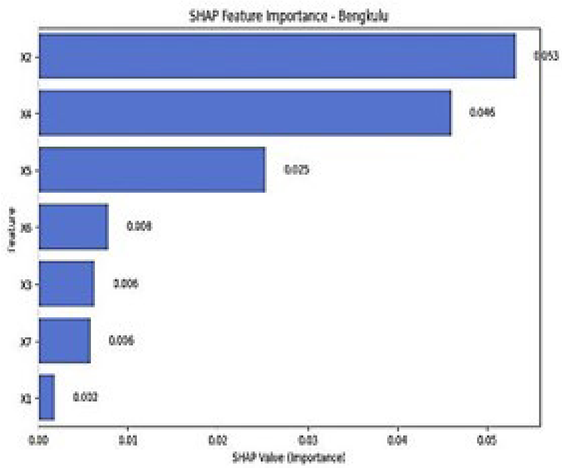
(a)



(b)

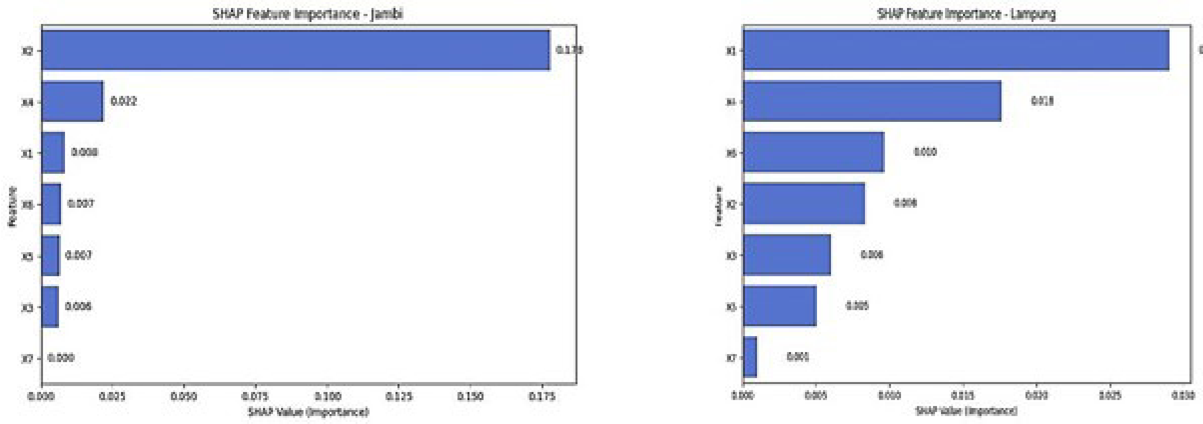


(c)



(d)

Figure 4 Traditional SHAP Plot by Province: (a) Riau (b) Kepulauan Riau (c) Bangka Belitung (d) Bengkulu



(a) (b)

Figure 5 Traditional SHAP plot by Province: (a) Jambi (b) Lampung

with adequate sanitation (X_5). For Jambi Province, the SHAP plot, shown in Figure 5 (see Appendices), indicates that the most important features are Underweight (X_2), Gross Regional Domestic Product per Capita (X_4), and Wasting (X_1).

In Kepulauan Riau Province, the three most important features are Underweight (X_2), Gross Regional Domestic Product per capita (X_4), and Households with adequate sanitation (X_5). For Lampung Province, Figure 5 shows that the three most important features are Wasting (X_1), Gross Regional Domestic Product per capita (X_4), and Households with adequate sanitation (X_5), For Riau Province, the three most important features are Gross Regional Domestic Product per Capita (X_4), Underweight (X_2), proportion of ever-married women aged 15–49 who gave birth to their first child (X_7). Overall, the dominant features across these provinces consistently include Gross Regional Domestic Product per capita (X_4), Underweight (X_2), and Households that have adequate sanitation (X_5). For West Sumatra Province, the three key features are Underweight (X_2), Households that have adequate sanitation (X_5), and Gross Regional Domestic Product per capita (X_4). For South Sumatra and North Sumatra Province, the three most important features are Underweight (X_2), Number of health worker (X_6), and the proportion of poor people (X_3).

For Therefore, the success of this research is benchmarked against model performance and model explainability criteria. First, the Random Forest model demonstrates the highest predictive accuracy among all tested models, including Decision Tree and Support Vector Machine, indicating superior performance in predicting stunting prevalence in Sumatra Province. This superiority is quantitatively reflected in higher accuracy values and lower prediction errors compared to the other models. Second, the explainability benchmark is achieved through the application of SHAP values, which successfully identify and explain the most influential features affecting stunting prevalence. The clear identification of key determinants, Underweight (X_2), Gross Regional Domestic Product per capita (X_4), and Households with adequate sanitation (X_5), demonstrates that the model is both interpretable and consistent with domain knowledge and policy relevance. Consequently, the benchmark for the success of this research lies in the Random Forest model's ability to achieve the highest predictive accuracy while simultaneously providing interpretable and policy-relevant explanations of the factors influencing stunting prevalence in Sumatra Province through SHAP analysis.

IV. CONCLUSIONS

The Random Forest model demonstrates the highest accuracy compared to the Decision Tree and Support Vector Machine models. In this study, the Random Forest model is used to predict stunting prevalence across provinces in Sumatra. The results incorporate SHAP values applied to stunting prevalence data in Sumatra Province to explain and interpret feature importance in the prediction of stunting.

The results show that Underweight (X_2) is the most influential factor for the Indonesian government in efforts to minimize stunting prevalence in Sumatra Province. In addition, the Gross Regional Domestic Product per capita (X_4) and the number of households with adequate sanitation (X_5) should also be considered by the Indonesian government to reduce the prevalence of Stunting in Sumatra Province.

A limitation of this study lies in the use of only seven independent variables and four model types. The restriction to a limited set of predictor variables and machine learning models may constrain predictive performance and reduce opportunities to explore more advanced or alternative modeling approaches. Future research is therefore encouraged to incorporate a broader range of predictor variables and to evaluate additional machine learning algorithms to further improve model performance and generalizability. Moreover, deeper exploration of the SHAP framework is recommended, particularly in its application to more diverse and complex machine learning models. Future studies may compare different SHAP variants, such as Kernel SHAP, Tree SHAP, and Deep SHAP, to assess their relative strengths, computational efficiency, and suitability for public health datasets.

Establishing clear benchmarks for SHAP visualization is also valuable, as such benchmarks can help users interpret feature contributions more effectively, reduce ambiguity, and enhance the practical usefulness of model explanations for policymakers and health practitioners. In addition, investigating advanced visualization techniques or integrating SHAP with complementary interpretability tools may provide richer insights into model behavior. These developments not only expand the methodological contribution of SHAP to classification problems but also create opportunities to identify novel patterns and actionable factors related to stunting and other public health outcomes.

AUTHOR CONTRIBUTIONS

Conceived and designed the analysis; Collected the data; Contributed data or analysis tools; Performed the analysis; Wrote the paper, A. A. P., N. F. S., K. A. N. and B. S.

DATA AVAILABILITY

The data that support the findings of this study are available in [Survei Kesehatan Indonesia Tahun 2023] at [https://www.badankebijakan.kemkes.go.id/hasil-ski-2023/]. These data were derived from the following resources available in the public domain: [https://www.badankebijakan.kemkes.go.id/hasil-ski-2023/].

REFERENCES

Arsenault, P.-D., Wang, S., & Patenaude, J.-M. (2025). A survey of explainable artificial intelligence (XAI) in financial time series forecasting. *ACM Computing Surveys*, *57*(10), 1–37. https://doi.org/10.1145/3729531

Asgedom, Y. S., Seifu, B. L., Mare, K. U., Asmare, Z. A., Asebe, H. A., Kase, B. F., Shibeshi, A. H., Tebeje, T. M., Sabo, K. G., & Fente, B. M. (2024). Levels of stunting associated factors among under-five children in Ethiopia: A multi-level ordinal logistic regression analysis. *Plos One*, *19*(1), e0296451. https://doi.org/10.1371/journal.pone.0296451

Ashari, R., Basyir, V., Afriwardi, A., Mayetti, M., Yusrawati, Y., & Desmawati, D. (2023). Factors Related to Stunting Incidence in Toddlers Aged 24-59 Months in the Working Area of Kambang Community Health Center, Pesisir Selatan District. *Contagion: Scientific Periodical Journal of Public Health and Coastal Health*, *5*(2), 530–549.

Asmare, A. A., Tegegne, A. S., Belay, D. B., & Agmas, Y. A. (2025). Coexisting predictors for undernutrition indices among under-five children in West Africa: Application of a multilevel multivariate ordinal logistic regression model. *BMC Nutrition*, *11*(1), 112. https://doi.org/10.1186/s40795-025-01099-x

Fadmi, F. R., Mulyani, S., Justin, W. O. S., & Riza, Y. (2025). Logistic Regression Analysis of Risk Factors for Stunting Among Toddlers Aged 24-59 Months in Southeast Sulawesi, Indonesia. *Health Dynamics*, *2*(2), 85–91. https://doi.org/10.33846/hd20206

Fahrani, D., Putri, A. E., & Pramana, S. (2025). *Combining survey and satellite data for spatial analysis of the prevalence of stunting in java in 2021*. *3302*(1), 050005. https://doi.org/10.1063/5.0262277

Falah, A. N., Andriyana, Y., Jaya, I., Tantular, B., & Maryadi, E. (2025). Expanded spatial Durbin model for analyzing stunting prevalence in Java Island. *Commun. Math. Biol. Neurosci.*, *2025*, Article-ID. https://doi.org/10.28919/cmbn/9217

Fedyk, M., & Ray, M. (2023). *How to leverage machine learning interpretability and explainability to generate hypotheses in cognitive psychology*. *45*(45).

Girma, B., Sasahu, L. D., & Rahman, A. (2025). Spatial distribution of stunting among breast feeding children in Sub-Sahara Africa. *PLoS One*, *20*(6), e0325812. https://doi.org/10.1371/journal.pone.0325812

Henninger, M., Debelak, R., Rothacher, Y., & Strobl, C. (2023). Interpretable machine learning for psychological research: Opportunities and pitfalls. *Psychological Methods*. https://doi.org/10.1037/met0000560

Houssein, E. H., Gamal, A. M., Younis, E. M., & Mohamed, E. (2025). Explainable artificial intelligence for medical imaging systems using deep learning: A comprehensive review. *Cluster Computing*, *28*(7), 469. https://doi.org/10.1007/s10586-025-05281-5

Juniarti, N., Alsharaydeh, E., Sari, C. W. M., Yani, D. I., & Hutton, A. (2025). Determinant factors influencing stunting prevention behaviors among working mothers in West Java Province, Indonesia: A cross-sectional study. *BMC Public Health*, *25*(1), 2719. https://doi.org/10.1186/s12889-025-24078-0

Kassie, G. A., & Asgedom, Y. S. (2025). Childhood stunting severity level and associated factors among under-five children in Tanzania: A multi-level ordinal logistic regression analysis using 2022 Tanzanian demographic and health survey. *BMC Pediatrics*, *25*(1), 129. https://doi.org/10.1186/s12887-025-05490-2

Mohsin, M. T., & Nasim, N. B. (2025). Explaining the Unexplainable: A Systematic Review of Explainable AI in Finance. *arXiv Preprint arXiv:2503.05966*. https://doi.org/10.48550/arXiv.2503.05966

Muhaimin, A., Ekacitta, P. C., & Ardiani, A. E. (2025). Spatial Autocorrelation Analysis of East Java Stunting Prevalence Cases in 2023. *Journal of Advances in Information and Industrial Technology*, *7*(1), 83–94.https://doi.org/10.52435/jaiit.v7i1.689

Purnamasari, I., Widiyati, F., & Sahli, M. (2022). Analisis Faktor yang Mempengaruhi Kejadian Stunting pada

Balita. *Jurnal Penelitian Dan Pengabdian Kepada Masyarakat UNSIQ*, *9*(1), 48–56. https://doi.org/10.32699/ppkm.v9i1.2342

Rehman, A., Lin, J. C., & Heldal, I. (2025). Enhancing Psychologists' Understanding Through Explainable Deep Learning Framework for ADHD Diagnosis. *Expert Systems*, *42*(2), e13788. https://doi.org/10.1111/exsy.13788

Rifada, M., Chamidah, N., Ningrum, R. A., & Muniroh, L. (2023). Stunting determinants among toddlers in Probolinggo district of Indonesia using parametric and nonparametric ordinal logistic regression models. *Commun. Math. Biol. Neurosci.*, *2023*, Article-ID. https://doi.org/10.28919/cmbn/6690

Shah, T., Shekokar, K., Barve, A., & Khandare, P. (2024). *An Analytical Review: Explainable AI for Decision Making in Finance Using Machine Learning*. 1–5. 10.1109/PICET60765.2024.10716075

Shifa, N., Saleh, M., Akbari, Y., & Al Maadeed, S. (2025). A review of explainable AI techniques and their evaluation in mammography for breast cancer screening. *Clinical Imaging*, *123*, 110492. https://doi.org/10.1016/j.clinimag.2025.110492

Sun, X., Liu, C., Wang, J., & Li, J. (2020). Assessing the extreme risk spillovers of international commodities on maritime markets: A GARCH-Copula-CoVaR approach. *International Review of Financial Analysis*, *68*, 101453.https://doi.org/10.1016/j.irfa.2020.101453

Uban, A.-S., Chulvi, B., & Rosso, P. (2022). Explainability of depression detection on social media: From deep learning models to psychological interpretations and multimodality. In *Early Detection of Mental Health Disorders by Social Media Monitoring: The First Five Years of the eRisk Project* (pp. 289–320). Springer. https://doi.org/10.1007/978-3-031-04431-1_13

Ullah, N., Guzmán-Aroca, F., Martínez-Álvarez, F., De Falco, I., & Sannino, G. (2025). A novel explainable AI framework for medical image classification integrating statistical, visual, and rule-based methods. *Medical Image Analysis*, 103665. https://doi.org/10.1016/j.media.2025.103665