MSTS Methods in Science and Technology Studies

futuristech PT. Teknologi Futuristik Indonesia

Date of publication November 22, 2025, date of current version November 22, 2025.
Digital Object Identifier 10.64539/msts.v1i1.2025.351

Article

Machine Learning-Based Early Prediction of Stunting Risk: A Comparative Study

Abdul Fadlil[1], Dikky Praseptian M[2], Muhammad Ma'ruf[3], Furizal[4,*]

[1] Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta 55191, Indonesia
[2] Department of Information Systems, STMIK PPKIA Tarakanita Rahmawati, Tarakan, North Kalimantan, Indonesia
[3] Sekolah Tinggi Ilmu Kesehatan ISFI Banjarmasin, Banjarmasin, Indonesia, 70123; maruf@stikes-isfi.ac.id
[4] Department of Research and Development, Peneliti Teknologi Teknik Indonesia, Yogyakarta 55281, Indonesia; furizal.id@gmail.com
* Correspondence

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Abstract: Stunting remains a critical nutritional issue among children, affecting growth and long-term human resource quality. Despite national programs and global targets for stunting reduction, early prediction of stunted children using data-driven methods remains limited. This study aims to evaluate and compare the performance of four supervised machine learning algorithms—Naïve Bayes, Multilayer Perceptron (MLP), Decision Tree (J48), and Support Vector Machine (SVM)—in predicting stunting using a dataset of 97 child records from three villages in East Kalimantan, Indonesia. Data were tested in both unnormalized and normalized forms and split into training and testing sets at 70%–30%, 80%–20%, and 90%–10% ratios. The results indicate that MLP and Decision Tree consistently achieved 100% accuracy across all splits and preprocessing conditions, while Naïve Bayes and SVM showed lower and more variable accuracy in certain cases. These findings suggest that MLP and Decision Tree are the most reliable methods for stunting prediction in small datasets, providing a practical approach for early identification and intervention. The study highlights the importance of algorithm selection and preprocessing in achieving optimal predictive performance in health-related datasets.

Keywords: Naïve Bayes; Multilayer Perceptron; Decision Tree; Support Vector Machine; Stunting Prediction.

1. Introduction

Stunting remains a severe nutritional problem among toddlers, characterized by shorter height compared to their peers, and is recognized as a critical public health issue in Indonesia [1]. Preventing stunting has become a national priority, reflected in Presidential Regulation No. 72 of 2022 [2] and the National Strategy for the Acceleration of Stunting Prevention 2018–2024 [3], as well as national stunting summits attended by governors, regents/mayors, and village heads to coordinate multisectoral efforts [4]. Despite these interventions, stunting prevalence in Indonesia has remained high over the past decade at around 37%, exceeding neighboring countries such as Thailand (16%), Vietnam (23%), and Myanmar (35%), ranking Indonesia fifth globally [5]. Although prevention programs provide nutrition education and food, early detection of toddlers at risk of stunting is crucial for targeted interventions [6].

Machine learning (ML) methods have shown potential for predicting stunting by analyzing nutritional and demographic data. Previous studies, however, reveal limitations. Hindratmo et al. (2021) achieved 97% accuracy using K-Nearest Neighbors (k-NN), but tested only a single method with a 90/10 train-test split [7]. Vega Herliansyah et al. (2021) obtained 64.02% accuracy with Naïve Bayes under the same data split [8]. Chilyabanyama et al. (2022) compared four classifiers, with Random Forest achieving 79% accuracy and Naïve Bayes 61.6%, but excluded artificial neural networks [9]. Similarly, Eva Darnila et al. (2022) reported 58.6% accuracy with Random Forest using two trial models [10], and M. Syauqi Haris et al. (2022) found Support Vector Regression (SVR) performed best among regression-based methods, but only regression algorithms were tested [11]. These findings indicate that while ML is promising for stunting prediction, prior research has often been limited by using a single method, insufficient comparison, or constrained data splits.

To address these gaps, this study evaluates and compares four machine learning algorithms—Naïve Bayes, Multilayer Perceptron (MLP), Decision Tree, and Support Vector Machine (SVM)—to determine the most effective approach for predicting stunting in toddlers. By including both probabilistic and neural network methods, this study provides a comprehensive evaluation and aims to identify a high-accuracy model that can support early intervention programs. The contributions of this research are:

a) comparative analysis of multiple supervised learning algorithms for stunting prediction;
b) inclusion of both traditional and neural network-based methods, and;
c) recommendations for practical application in early detection and targeted interventions.

The remainder of the paper is structured as follows: Section 2 presents the basic concepts of the machine learning methods used; Section 3 describes the research methodology; Section 4 presents the results and discussion; and Section 5 concludes the study and outlines future research directions.

2. Basic Concepts

2.1. Naïve Bayes

Naïve Bayes is a simple yet effective supervised learning algorithm widely used for classification tasks [12], [13]. It is based on Bayes' theorem in statistics and adopts a strong assumption: the feature values are conditionally independent given the class label [14]. Despite this "naïve" assumption, the algorithm often performs well in practice.

The Naïve Bayes classifier computes the probability of a data instance belonging to a particular class using the likelihood of its features. Formally, let c_i represent a class label, C the set of all possible classes, t_f a single feature, and F the total number of features. The likelihood of a class given the features is calculated as:

$$likelihood(c_i) = P(c_i)\prod_{f=1}^{F} P(t_f|c_i) \quad (1)$$

The assignment probability for each class is then obtained by normalizing the likelihoods across all classes:

$$Passignment(c_i) = \frac{likelihood(c_i)}{\sum_{c_i \in C} likelihood(c_j)} \quad (2)$$

Finally, the predicted class is the one with the highest assignment probability:

$$c_i = \underset{c_i \in C}{\arg max}\ Passignment(c_i) \quad (3)$$

Advantages of Naive Bayesian: a. Handles quantitative and discrete data. b. It requires very little training data to estimate the parameters (mean and variable variance) required for classification. c. Handling missing value by ignoring agency during estimation of likelihood calculations. d. Fast and space efficiency. e. Robust to irrelevant attributes Naive Bayesian losses: 1) Not applicable if the conditional probability is zero, if it is zero then the predicted probability will be zero as well, 2) Assuming an independent variable. Assumption of independent variables.

Advantages of Naïve Bayes:

a) Can handle both quantitative and discrete data.
b) Requires relatively little training data to estimate the necessary parameters (mean and variance).
c) Can handle missing values by ignoring them during likelihood estimation.
d) Computationally fast and space-efficient.
e) Robust to irrelevant features.

Limitations of Naïve Bayes:

a) If a conditional probability is zero, the predicted probability becomes zero as well (can be mitigated with smoothing techniques).
b) Assumes independence among features, which may not hold in practice, potentially reducing accuracy [15].

2.2. Multilayer Perceptron

Multilayer Perceptron (MLP) is a type of Artificial Neural Network (ANN) commonly used for supervised learning tasks [16]. An MLP typically consists of an input layer, one or more hidden layers, and an output layer [17], [18]. The network is trained through two main phases: feedforward and backpropagation. In the feedforward phase, the input layer receives the raw input values and passes them to the hidden layer, where each neuron computes a weighted sum of its inputs, adds a bias, and applies an activation function, as described by:

$$o_j = \sigma\left(\sum_{k=1}^{K} x_k w_{k,j} + \beta_j\right) \quad (4)$$

where x_k is the k-th input, $w_{k,j}$ is the weight connecting input k to hidden neuron j, β_j is the bias of neuron j, and σ is the activation function. The outputs of the hidden layer are then passed to the output layer, which computes its outputs using:

$$v_i = \sigma\left(\sum_{j=1}^{J} o_j u_{j,i} + \gamma_i\right) = = \sigma\left(\sum_{j=1}^{J} \sigma\left(\sum_{k=1}^{K} x_k w_{k,j} + \beta_j\right) u_{j,i} + \gamma_i\right) \quad (5)$$

where $u_{j,i}$ is the weight connecting hidden neuron j to output neuron i, and γ_i is the bias of the output neuron. In the backpropagation phase, the network updates its weights and biases to minimize the difference between predicted and target outputs [19]. MLPs are generally fully connected, meaning every neuron in one layer is connected to every neuron in the next layer. By combining multiple non-linear functions across layers and optimizing the parameters through gradient-based methods, MLPs can model complex relationships that a single perceptron cannot, making them a powerful tool for both classification and regression tasks.

2.3. Decision Tree

A Decision Tree is a tree-like structure used to model decisions and their possible outcomes, including resource costs, utilities, and risks [20], [21]. The process begins at the root node, where a feature is evaluated, and a branch is chosen based on the feature's value. This process continues sequentially down the tree until a leaf node is reached, which represents the target class or the final decision. Decision Tree 4.5, for instance, has the advantage of being able to handle data in various formats.

Decision trees are a form of inductive learning, where the tree is constructed based on the training data [22], [23]. A well-known variant of the decision tree algorithm is ID3 [24], - [26]. The fundamental assumption is that if the attributes provide sufficient information, the decision tree can correctly classify all instances in the training dataset. To generalize to unseen instances, the tree must capture the underlying relationships between attributes and class values.

Conceptually, a decision tree consists of nodes, branches, and leaves. Each node represents a decision based on a particular attribute, each branch represents the outcome of that decision, and the path from the root to a leaf encodes a classification rule. This structure allows the algorithm to make decisions in a systematic, interpretable way, where conditional statements guide the decision-making process and lead to the desired outcome.

2.4. Support Vector Machine

Support Vector Machine (SVM) is a supervised learning algorithm widely used for classification and pattern recognition [27], [28]. SVM is derived from statistical learning theory and is designed to provide high accuracy compared to other machine learning methods [29]. SVM works by mapping input data into a high-dimensional feature space and finding an optimal hyperplane that separates different classes. The algorithm identifies this hyperplane using support vectors, which are the data points closest to the decision boundary, and maximizes the margin, the distance between the hyperplane and the nearest support vectors. This approach ensures that the classifier generalizes well to unseen data. The theoretical foundation of SVM has evolved since the 1960s and was formally introduced by Vapnik, Boser, and Guyon in 1992.

SVM can handle both linear and non-linear classification problems through the use of kernel functions and can also be applied to regression tasks (Support Vector Regression). Compared to other classification techniques, SVM offers a mathematically rigorous framework and a clear geometric interpretation, making it a robust and widely used method in machine learning [30].

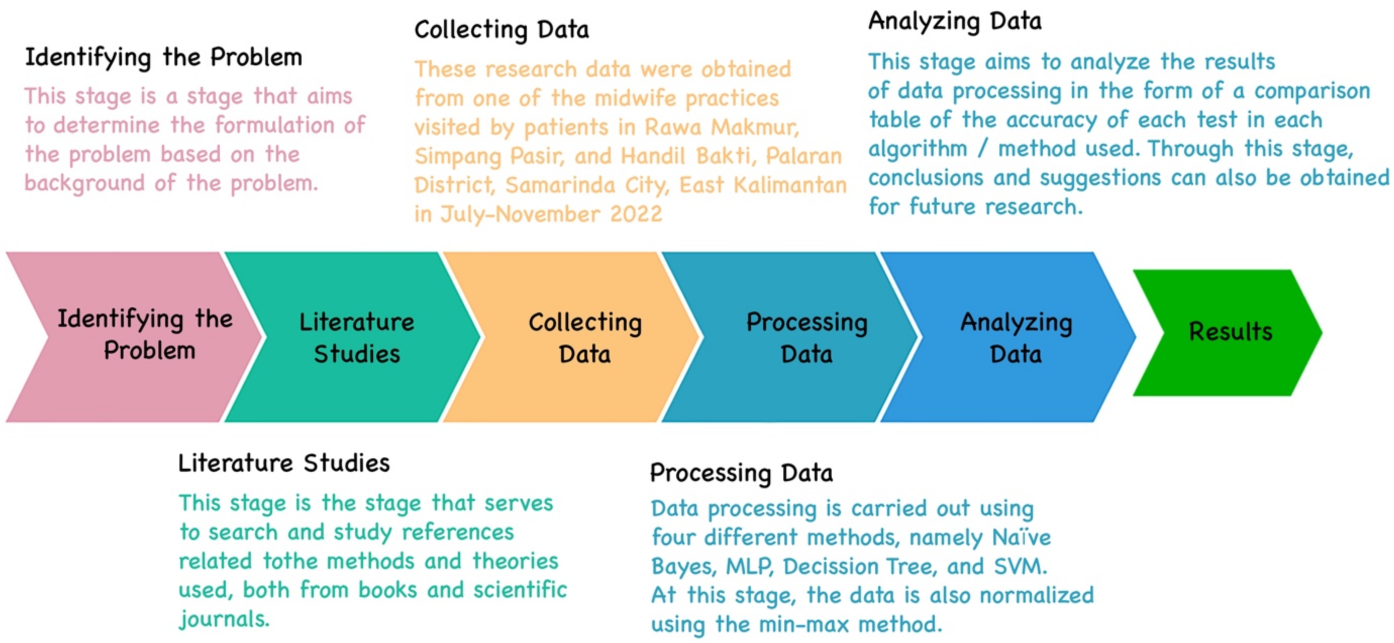

Figure 1. Stages of research.

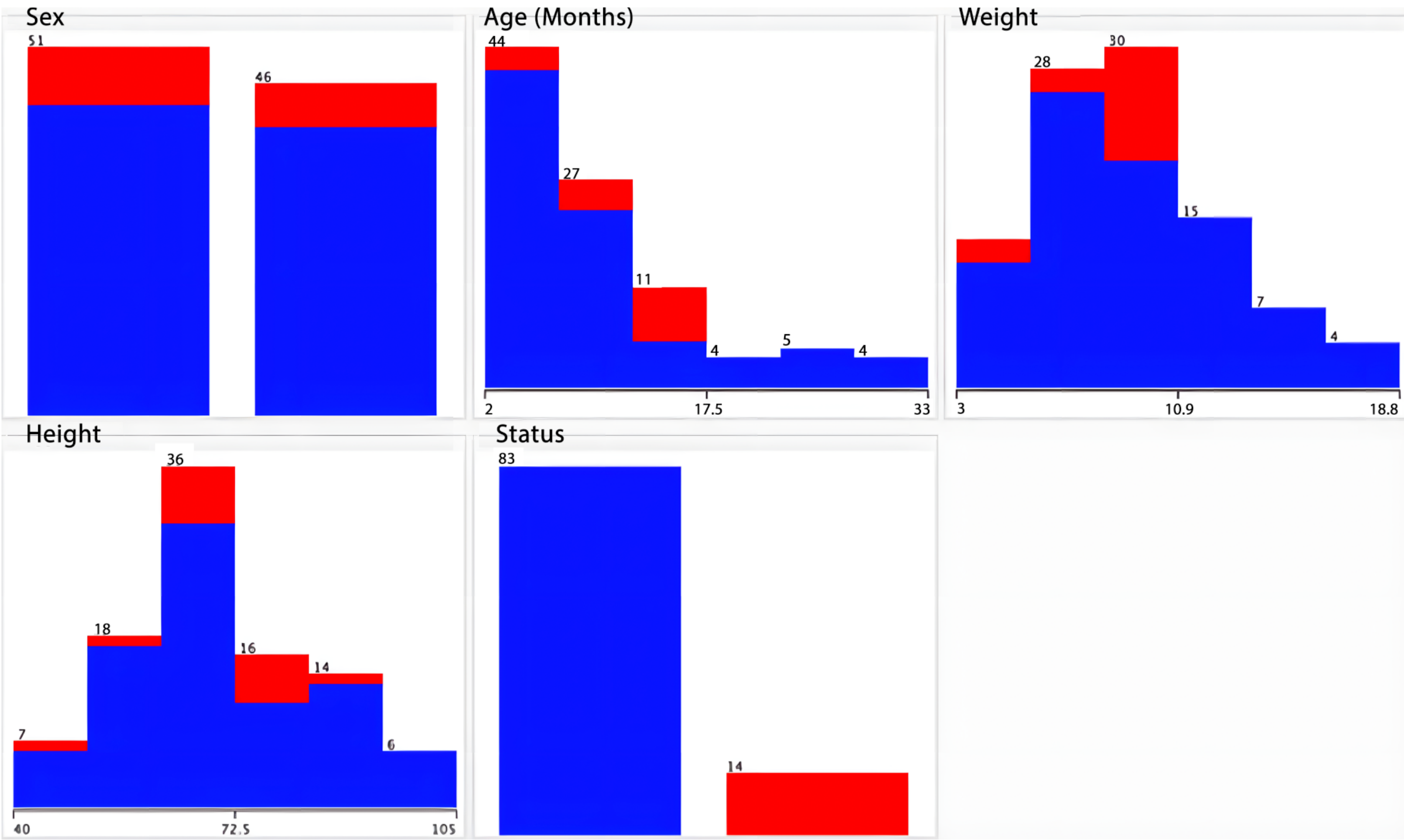

Figure 2. dataset visualization.

3. Research Method

The stages of research conducted in this study are illustrated in Figure 1. This study consists of seven main stages: problem identification, literature review, data collection, data processing, data analysis, and result interpretation. The expected outcome is a comparison of four machine learning methods—Naïve Bayes, Multilayer Perceptron (MLP), Decision Tree (J48), and Support Vector Machine (SVM)—in predicting stunting in children, with the objective of identifying the method with the highest accuracy.

3.1. Identifying the Problem

Problem identification focuses on determining the most effective machine learning method for predicting stunting with high accuracy while minimizing overfitting. Considering stunting prevention is a national program in Indonesia and part of the Global Nutrition Target for 2025, accurate prediction can support early interventions and improve human resource quality. The presence of multiple machine learning algorithms necessitates systematic testing and comparison to determine the most suitable algorithm for this purpose.

3.2. Literature Studies

After identifying the problem, relevant literature was reviewed from books and journals to understand previous research and methodologies in stunting prediction using machine learning. The literature indicates varying accuracy across methods and highlights the need for a comprehensive comparison of multiple algorithms, including probabilistic, tree-based, and neural network approaches.

3.3. Collecting Data

Data for this study consists of 97 records of toddlers' nutritional status, collected from midwife practices between July and November 2022. The data were sourced from three villages—Rawa Makmur, Simpang Pasir, and Handil Bakti—located in Palaran District, Samarinda City, East Kalimantan, Indonesia. The dataset includes four attributes: gender (binomial), age, weight, and height (continuous), with the class label being nutritional status, categorized as malnutrition or normal nutrition. The dataset was visualized in Figure 2.

3.4. Processing Data

Data preprocessing was performed using Weka Tools version 3.8.6. The Naïve Bayes, MLP, and Decision Tree (J48) methods are available by default, while SVM requires the additional library function.libSVM. The preprocessing steps include normalization to equalize

Table 1. Split ratios.

No	Training Data (%)	Testing Data (%)
1	90	10
2	80	20
3	70	30

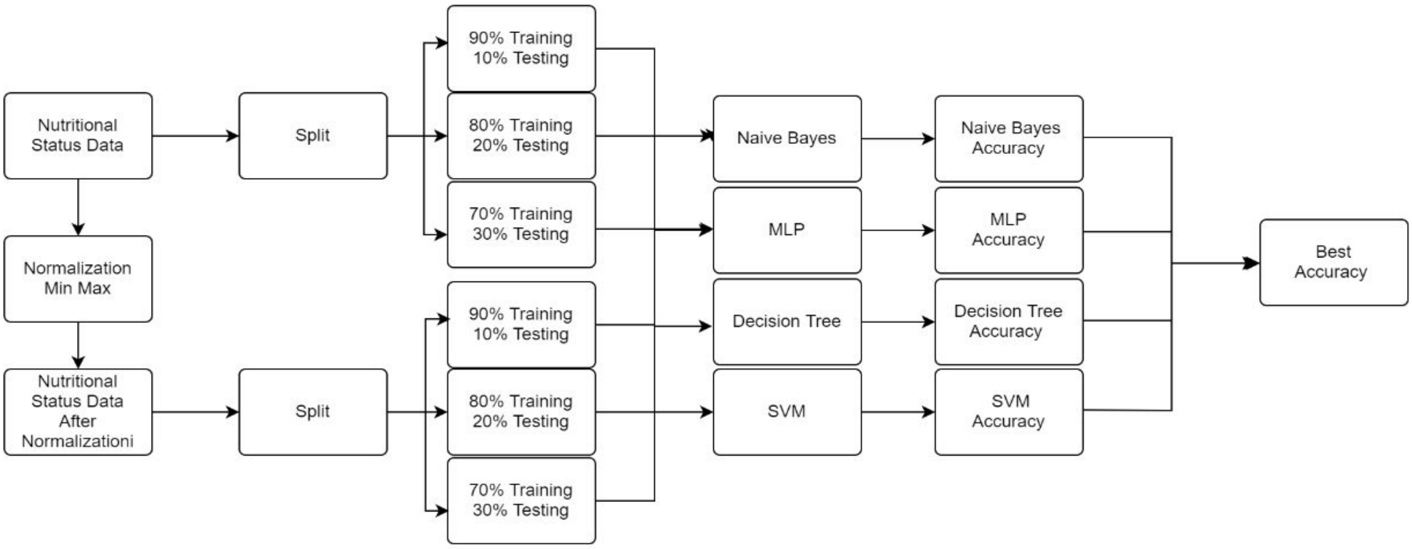

Figure 3. Principles of analysis.

Table 2. Test comparison results of each method.

No	Algorithm Name	Algorithm Name in Weka App	Data Training (%)	Data Testing (%)	Accuracy (%)	
					Unnormalized	Normalized
1	Naïve bayes	bayes.NaïveBayes	70	30	89.65	89.65
			80	20	89.47	89.47
			90	10	100	100
2	MLP	functions.Multilayer-Perceptron	70	30	100	100
			80	20	100	100
			90	10	100	100
3	Decision Tree	trees.J48	70	30	100	100
			80	20	100	100
			90	10	100	100
4	SVM	functions.libSVM	70	30	93.10	89.65
			80	20	84.21	89.47
			90	10	100	100

attribute ranges, preventing excessively large or small values from biasing statistical analysis. The min-max normalization method was used, as it has been shown to produce optimal results in previous studies [31]. The min-max normalization formula is given in Equation 6.

$$x_{new} = \frac{x_{old} - x_{min}}{x_{max} - x_{min}} \quad (6)$$

Where x_{new} is the normalized value, x_{old} is the original value, and x_{max} and x_{min} are the maximum and minimum values of the dataset, respectively.

Besides that, the dataset was divided into training and testing sets to evaluate model performance. Three split ratios were used, as shown in Table 1, and the effect of normalization on model accuracy was also assessed.

3.5. Analyzing Data

Based on the analysis principles stated in Figure 3, the data is tested in two types: the original data (unnormalized) and the normalized data. These datasets are evaluated using four machine learning methods—Naïve Bayes, Multilayer Perceptron (MLP), Decision Tree (J48), and Support Vector Machine (SVM)—implemented in Weka Tools version 3.8.6 to provide a comprehensive comparison of their performance.

3. Results and Discussion

Based on the comparison in Table 2, the highest accuracy was achieved by the Multilayer Perceptron (MLP) and Decision Tree (J48) algorithms, which consistently reached 100% accuracy across all training-testing splits. In contrast, Naïve Bayes and Support Vector Machine (SVM) exhibited lower accuracy in certain test cases, particularly with training-testing splits of 70%–30% and 80%–20%.

The heatmaps in Figure 4 summarize the minimum, mean, and maximum accuracy for each method across different splits and for both unnormalized and normalized datasets. From Figure 4:

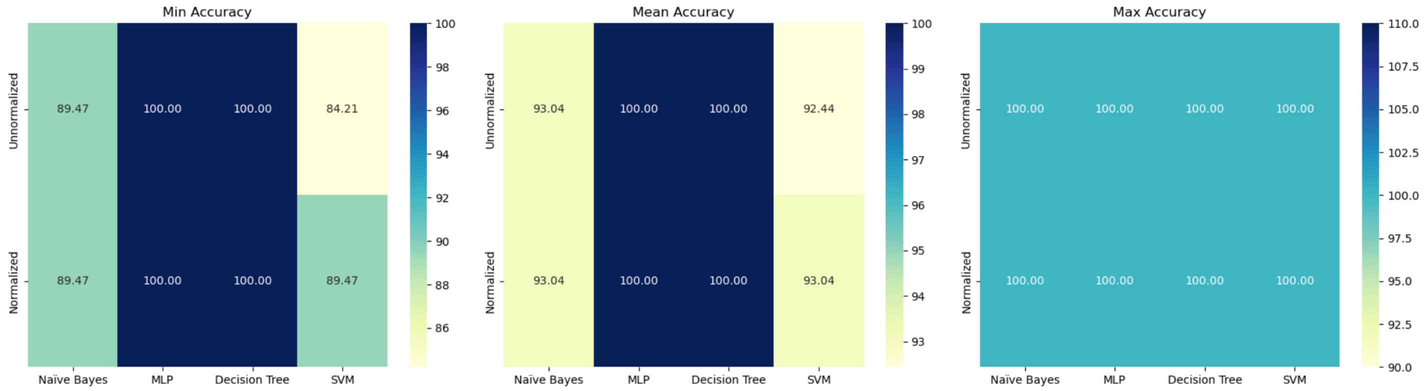

Figure 4. Heatmap of minimum, mean, and maximum accuracy for four machine learning methods (Naïve Bayes, MLP, Decision Tree, SVM) using unnormalized and normalized stunting datasets.

a) MLP and Decision Tree (J48) show 100% accuracy for all conditions.
b) Naïve Bayes has minimum accuracy 89.47% and mean accuracy 93.04% for both unnormalized and normalized datasets.
c) SVM has minimum accuracy 84.21% for unnormalized data and 89.47% for normalized data. The mean accuracy is 92.44% for unnormalized and 93.04% for normalized datasets.

These results indicate that MLP and Decision Tree (J48) consistently achieved the highest accuracy, while Naïve Bayes and SVM had lower accuracy in certain splits. The comparison between unnormalized and normalized data shows minor differences for some methods, particularly SVM.

4. Conclusion

Based on the comparison of machine learning methods for predicting stunting in children, Multilayer Perceptron (MLP) and Decision Tree (J48) consistently achieved the highest accuracy of 100% across all training-testing splits and data pre-processing conditions, demonstrating stable performance. Naïve Bayes and Support Vector Machine (SVM) showed lower accuracy in some splits, with minimum values of 89.47% and 84.21%, respectively, indicating some sensitivity to training data proportion and normalization. Overall, for the dataset used in this study, MLP and Decision Tree (J48) are the most reliable methods for stunting prediction, while Naïve Bayes and SVM require careful consideration of data pre-processing to achieve optimal results.

5. Declarations

5.1. Author Contributions

Abdul Fadlil: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation; **Dikky Praseptian M:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - Original Draft; **Muhammad Ma'ruf:** Formal analysis, Resources, Data Curation; **Furizal:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - Original Draft.

5.2. Institutional Review Board Statement

Not applicable.

5.3. Informed Consent Statement

Not applicable.

5.4. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

5.5. Acknowledgment

Not applicable.

5.6. Conflicts of Interest

The authors declare no conflicts of interest.

6. References

[1] T. Beal, A. Tumilowicz, A. Sutrisna, D. Izwardy, and L. M. Neufeld, "A review of child stunting determinants in Indonesia," *Matern Child Nutr*, vol. 14, no. 4, pp. 1–10, 2018, https://doi.org/10.1111/mcn.12617.

[2] Indonesian Government, "Pepres No 72 Tahun 2021," *Indonesian Government*, no. 1, p. 23, 2021.

[3] E. Satriawan, "Strategi Nasional Percepatan Pencegahan Stunting 2018-2024 (National Strategy for Accelerating Stunting Prevention 2018-2024)," *Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K) Sekretariat Wakil Presiden Republik Indonesia*, no. November, pp. 1–32, 2018.

[4] K. Riskerdas, "Hasil Utama Riset Kesehatan Dasar (RISKEDAS)," *J Phys A Math Theor*, vol. 8, no. 44, pp. 1–200, 2018.

[5] D. A. Sulistyo, Y. S. Putra, and S. Y. Riska, "Metode Agile dalam Pengembangan Sistem Prediksi Prevalensi Stunting di Indonesia," *Network Engineering Research Operation*, vol. 5, no. 2, p. 74, Oct. 2020, https://doi.org/10.21107/nero.v5i2.160.

[6] K. Rahmadhita, "Permasalahan Stunting dan Pencegahannya," *Jurnal Ilmiah Kesehatan Sandi Husada*, vol. 11, no. 1, pp. 225–229, 2020, https://doi.org/10.35816/jiskh.v11i1.253.

[7] H. H. Sutarno, R. Latuconsina, and A. Dinimaharawati, "Prediksi Stunting pada Balita dengan Menggunakan Algoriitma Klasifikasi K-Nearest Neighbors," *e-Proceeding of Engineering*, vol. 8, no. 5, pp. 6657–6661, 2021, Accessed: Nov. 22, 2025. [Online]. Available: https://repositori.telkomuniversity.ac.id/pustaka/files/171048/jurnal_eproc/prediksi-stunting-pada-balita-dengan-menggunakan-algoritma-klasifikasi-k-nearest-neighbors.pdf

[8] V. Herliansyah, R. Latuconsina, and A. Dinimaharawati, "Prediksi Stunting pada Balita dengan menggunakan Algoritma Klasifikasi Naive Bayes," *e-Proceeding of Engineering*, vol. 8, no. 5, pp. 6642–6649, 2021, Accessed: Nov. 22, 2025. [Online]. Available: https://repository.telkomuniversity.ac.id/pustaka/files/170976/jurnal_eproc/prediksi-stunting-pada-balita-dengan-menggunakan-algoritma-klasifikasi-na-ve-bayes.pdf

[9] O. N. Chilyabanyama *et al.*, "Performance of Machine Learning Classifiers in Classifying Stunting among Under-Five Children in Zambia," *Children*, vol. 9, no. 7, 2022, https://doi.org/10.3390/children9071082.

[10] E. Darnila, M. Maryana, K. Mawardi, M. Sinambela, and I. Pahendra, "Supervised models to predict the Stunting in East Aceh," *International Journal of Engineering, Science and Information Technology*, vol. 2, no. 3, pp. 33–39, 2022, https://doi.org/10.52088/ijesty.v2i3.280.

[11] M. S. Haris, A. N. Khudori, and W. T. Kusuma, "Perbandingan Metode Supervised Machine Learning untuk Prediksi Prevalensi Stunting di Provisi Jawa Timur," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 7, p. 1571, Dec. 2022, https://doi.org/10.25126/jtiik.2022976744.

[12] M. A. Yulianto and A. Fadlil, "Wood Type Identification System using Naive Bayes Classification," *Control Systems and Optimization Letters*, vol. 1, no. 3, pp. 139–143, Sep. 2023, https://doi.org/10.59247/csol.v1i3.52.

[13] A. Yudhana, D. Sulistyo, and I. Mufandi, "GIS-based and Naïve Bayes for nitrogen soil mapping in Lendah, Indonesia," *Sens Biosensing Res*, vol. 33, no. 1, pp. 1–13, Aug. 2021, https://doi.org/10.1016/j.sbsr.2021.100435.

[14] M. F. Nugraha and S. B. Rahayu, "Penerapan Naïve Bayes dalam Mengklasifikasi Calon Penerima Bantuan Pangan Non Tunai di Desa Nanjung Mekar," *INTERNAL (Information System Journal)*, vol. 5, no. 2, pp. 137–146, Dec. 2022, https://doi.org/10.32627/internal.v5i2.634.

[15] A. Y. P. Yusuf and R. Sari, "Implementasi Algoritma Naïve Bayes untuk Klasifikasi Pemahaman Program MBKM bagi Mahasiswa," *Journal of Informatic and Information Security*, vol. 3, no. 2, pp. 171–180, Dec. 2022, https://doi.org/10.31599/40dppk38.

[16] A. K. S. Lenson and G. Airlangga, "Comparative Analysis of MLP , CNN , and RNN Models in Automatic Speech Recognition : Dissecting Performance Metric," *Buletin Ilmiah Sarjana Teknik Elektro*, vol. 5, no. 4, pp. 576–583, 2023, https://doi.org/10.12928/biste.v5i4.9668.

[17] J. Naskath, G. Sivakamasundari, and A. A. S. Begum, "A Study on Different Deep Learning Algorithms Used in Deep Neural Nets: MLP SOM and DBN," *Wirel Pers Commun*, vol. 128, no. 4, pp. 2913–2936, Feb. 2023, https://doi.org/10.1007/s11277-022-10079-4.

[18] J. Zhang, C. Li, Y. Yin, J. Zhang, and M. Grzegorzek, "Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual

transformer," *Artif Intell Rev*, vol. 56, no. 2, pp. 1013–1070, Feb. 2023, https://doi.org/10.1007/s10462-022-10192-7.

[19] Y. Miftahuddin and M. M. Faturrahman, "Penerapan Data Standardization dan Multilayer Perceptron pada Identifikasi Website Phishing," *MIND (Multimedia Artificial Intelligent Networking Database) Journal*, vol. 7, no. 2, pp. 111–123, 2022, https://doi.org/10.26760/mindjournal.v7i2.111-123.

[20] G. Luo, M. A. Arshad, and G. Luo, "Decision Trees for Strategic Choice of Augmenting Management Intuition with Machine Learning," *Symmetry (Basel)*, vol. 17, no. 7, p. 976, Jun. 2025, https://doi.org/10.3390/sym17070976.

[21] M. Enayati, O. Bozorg-Haddad, M. Pourgholam-Amiji, B. Zolghadr-Asli, and M. Tahmasebi Nasab, "Decision Tree (DT): A Valuable Tool for Water Resources Engineering," 2022, pp. 201–223. https://doi.org/10.1007/978-981-19-2519-1_10.

[22] R. K. Patra, A. Mahendar, and G. Madhukar, "Inductive Learning Including Decision Tree and Rule Induction Learning," in *Data Mining and Machine Learning Applications*, Wiley, 2022, pp. 209–234. https://doi.org/10.1002/9781119792529.ch9.

[23] D. Mihai and M. Mocanu, "Processing GIS Data Using Decision Trees and an Inductive Learning Method," *Int J Mach Learn Comput*, vol. 11, no. 6, pp. 393–398, Nov. 2021, https://doi.org/10.18178/ijmlc.2021.11.6.1067.

[24] M. Qiu, "Path Planning Algorithm and ID3 Decision Tree Model Application of Scenic Intelligent Navigation System," *Procedia Comput Sci*, vol. 247, pp. 1187–1196, 2024, https://doi.org/10.1016/j.procs.2024.10.143.

[25] C. Liu, J. Lai, B. Lin, and D. Miao, "An improved ID3 algorithm based on variable precision neighborhood rough sets," *Applied Intelligence*, vol. 53, no. 20, pp. 23641–23654, Oct. 2023, https://doi.org/10.1007/s10489-023-04779-y.

[26] S. Kraidech and K. Jearanaitanakij, "Improving ID3 Algorithm by Combining Values from Equally Important Attributes," in *2017 21st International Computer Science and Engineering Conference (ICSEC)*, IEEE, Nov. 2017, pp. 1–5. https://doi.org/10.1109/ICSEC.2017.8443862.

[27] P. Chhajer, M. Shah, and A. Kshirsagar, "The applications of artificial neural networks, support vector machines, and long–short term memory for stock market prediction," *Decision Analytics Journal*, vol. 2, p. 100015, Mar. 2022, https://doi.org/10.1016/j.dajour.2021.100015.

[28] A. M. Elshewey, M. Y. Shams, N. El-Rashidy, A. M. Elhady, S. M. Shohieb, and Z. Tarek, "Bayesian Optimization with Support Vector Machine Model for Parkinson Disease Classification," *Sensors*, vol. 23, no. 4, p. 2085, Feb. 2023, https://doi.org/10.3390/s23042085.

[29] Y. Afrillia, L. Rosnita, and D. Siska, "Analisis Sentimen Ciutan Twitter Terkait Penerapan Permendikbudristek Nomor 30 Tahun 2021 Menggunakan TextBlob dan Support Vector Machine," *G-Tech: Jurnal Teknologi Terapan*, vol. 6, no. 2, pp. 387–394, Oct. 2022, https://doi.org/10.33379/gtech.v6i2.1778.

[30] F. S. Jumeilah, "Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 1, no. 1, pp. 19–25, 2017, https://doi.org/10.29207/resti.v1i1.11.

[31] D. A. Nasution, H. H. Khotimah, and N. Chamidah, "Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritma K-NN," *Computer Engineering, Science and System Journal*, vol. 4, no. 1, p. 78, 2019, https://doi.org/10.24114/cess.v4i1.11458.