JOIN (Jurnal Online Informatika)
p-ISSN: 2528-1682, e-ISSN: 2527-9165
Volume 10 Number 1 | June 2025: 53-65
DOI: 10.15575/join.v10i1.1495

Modality-based Modeling with Data Balancing and Dimensionality Reduction for Early Stunting Detection

Yohanes Setiawan[1], Moh Hamim Zajuli Al Faroby[2], Mochamad Nizar Palefi Ma'ady[3], I Made Wisnu Adi Sanjaya[2], Cisa Valentino Cahya Ramadhani[1]
[1]Department of Information Technology, Telkom University, Surabaya Campus, Surabaya, Indonesia
[2]Department of Data Science, Telkom University, Surabaya Campus, Surabaya, Indonesia
[3]Department of Information Systems, Telkom University, Surabaya Campus, Surabaya, Indonesia

Article Info

Article history:

Received October 30, 2024
Revised November 15, 2024
Accepted February 01, 2025
Published April 01, 2025

Keywords:

Data Balancing
Dimensionality Reduction
Multimodal
Stunting
Unimodal

ABSTRACT

In Indonesia, the stunting rate has reached 36%, significantly higher than the World Health Organization's (WHO) standard of 20%. This high prevalence underscores the urgent need for effective early detection methods. Traditional data mining approaches for stunting detection have primarily focused on unimodal data, either tabular or image data alone, limiting the comprehensiveness and accuracy of the detection models. Modality-based modeling, which integrates image and tabular data, can provide a more holistic view and improve detection accuracy. This research aims to analyze modality-based modeling for the early detection of stunting. Two modalities, unimodal and multimodal, are used in this study. The main contributions of this research are the development of a comprehensive framework for modality-based analysis, the application of advanced data preprocessing techniques, and the comparison of various machine learning algorithms to identify the best model for stunting detection. The dataset, comprising images and tabular data, is sourced from Posyandu in Sidoarjo, Indonesia. Image data undergoes preprocessing, including background segmentation and feature extraction using the Gray Level Co-occurrence Matrix (GLCM), while tabular data is processed through categorical encoding. The Synthetic Minority Oversampling Technique (SMOTE) addresses class imbalance, and Principal Component Analysis (PCA) is used for dimensionality reduction. Unimodal modeling uses tabular or image data alone, while multimodal modeling combines both before classification. The study achieves the best F1 scores of 0.96, 0.91, and 0.90 for tabular-only, image-only, and image-tabular modalities, respectively, demonstrating the effectiveness of data balancing and dimensionality reduction techniques.

Corresponding Author:

Moh Hamim Zajuli Al Faroby,
Department of Data Science, Telkom University, Surabaya Campus, Surabaya, Indonesia
Jl. Ketintang 156, Surabaya, Indonesia. 60231
Email: alfarobymhz@telkomuniversity.ac.id

1. INTRODUCTION

In Indonesia, the rate of stunting has reached 36%, which is higher than the 20% standard by the World Health Organization (WHO), but it is still lower than in India and Nepal, at 39% and 37%, respectively. [1]. Besides that, Indonesia is also known to have a high prevalence of stunted children in 2021 [2]This shows that Indonesia still must struggle to fight stunting. Stunting is chronic malnutrition resulting from inadequate nutritional intake during the first thousand days of a child's life, disrupting the quality of human resources. [3]. As the growth of information technology grows in daily life, a system that can be used to detect potential stunting early can be designed. Thus, the negative impact of stunting can be minimized as awareness of reducing stunted children is increased.

Researchers have already implemented information technology to detect stunting earlier, especially in data mining. Yuliana used Logistics Regression to determine whether a toddler has malnutrition status to detect stunting earlier as statistical analysis has been implemented [4]. Then, Gustriansyah detected stunting earlier using varied machine learning models to be compared, e.g., k-Nearest Neighbor, Naïve Bayes, Linear Discriminant Analysis, Decision Tree, Random Forest, and Support Vector Machine [5]. The more advanced analysis comes from Ningrum, which utilized feature engineering analysis besides using machine learning modeling [6]. However, data mining researchers of stunting detection focus only on tabular form analysis [7], leading to unimodal analysis. Researchers rarely use another modality based on data, namely multimodal analysis, to predict stunting, even if multimodal analysis is widely used for other applications.

Modality-based data mining analysis is used nowadays for medical analysis. This consists of two models: unimodal and multimodal analysis. Unimodal only provides a single piece of information about the dataset (tabular or image only), while multimodal gives a more comprehensive framework flexibility with multiple details about the dataset. [8]. Several researchers utilize multiple images as features for classification in biomedical research. [9], [10], which leads to deeper analysis rather than just a single image as the input of deep learning algorithms. Others have implemented images and non-images as multiple features for novel frameworks in medical data mining, including a combination of image (2D/3D/4D) and clinical data (tabular), and a combination of image (2D/3D/4D) and text data (natural language processing) [11]-[13]. These lead to a novel flexible framework in data mining and machine learning.

This paper proposes modality-based data mining modeling to detect stunting earlier. The first contribution is proposing frameworks based on modality-based analysis, consisting of unimodal analysis, tabular-only and image-only analysis, and multimodal analysis, which provides image-tabular analysis in the study case of stunting using primary data. The second contribution is proposing further data analysis by advanced pre-processing data, that is, utilization of data balancing by SMOTE for handling imbalanced datasets, the implementation of Principle Component Analysis (PCA) for dimensionality reduction, and the use of Grey Level Co-occurrence Matrix (GLCM) as the feature extraction for image dataset. Furthermore, this research uses various machine learning algorithms to be selected, including k-nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), Light Gradient Boosting Machine (LGBM), Catboost (CB), Adaboost (AB), Bagging Classifier (BC), and Multilayer Perceptron (MLP). The combination of pre-processing and various machine learning has been compared to find the best model for each modality-based analysis.

2. METHOD

This research examines stunting data through several steps: data gathering from qualitative and quantitative studies on stunting detection using machine learning techniques. The study utilizes the potential of a stunting detection dataset from two modalities. Figure 1 illustrates the methodology proposed for stunting detection, and the explanation follows the proposed method.

2.1. Data Collection

We collected primary data from qualitative research for image data gathering conducted at a health center in Sidoarjo City, Indonesia. In contrast, tabular data is collected from the prevalence of stunting tables in 2024. The initial data collected is 200, including 100 images and 100 tabular data. Then, each modality consists of 16 data, indicated as potentially stunted child data, and 84 as non-potentially stunted child data. Then, after the SMOTE data balancing, it becomes 84 stunting data for image and tabular. Hence, the total data is 336 data.

2.2. *Tabular and Image Data Preprocessing*

For the tabular data preprocessing, we assigned related image names to the tabular data and checked for any missing values. The tabular data takes three variables: age, gender, and weight. In the image-tabular modality, the tabular only takes two variables, age and gender, where the existence of the image replaces weight to predict whether there is potential for stunting or not, according to the KIA Book. We then assign related names to the tabular data to ensure no duplicated rows or children. For the image data preprocessing, we applied cropped images from their backgrounds so that the model could recognize the actual objects of the toddlers. Feature extraction in this study was conducted using the gray-level co-occurrence matrix (GLCM). This method captures textural characteristics within image data by evaluating the spatial relationships between pixel intensities. GLCM provides key textural features, including contrast, correlation, energy, and homogeneity, which are instrumental in analyzing images associated with stunting. The attributes selected for the tabular modality include sex, age, and weight, while the image data directly represents the imaging modality [14].

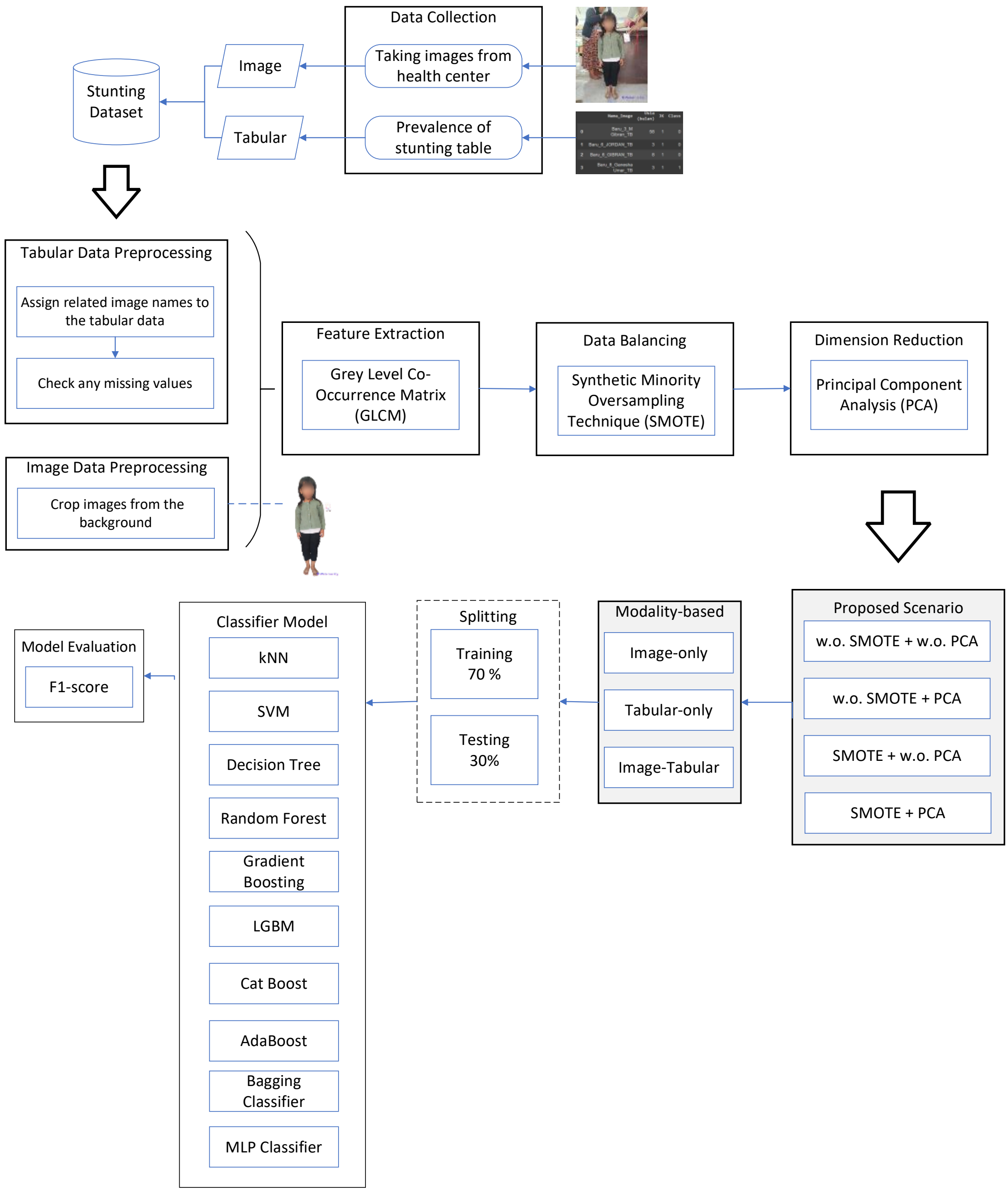


Figure 1. Proposed Methodology

2.3. Data Balancing and Dimensionality Reduction

The data have an imbalanced class, with significantly more samples in the non-stunted child class than in the stunting class. Methods such as random under-sampling, random oversampling, and SMOTE are applied to address this imbalance. After the feature extraction, data balancing is implemented through the Synthetic Minority Oversampling Technique (SMOTE). SMOTE addresses class imbalance by generating synthetic samples for the minority class, thereby reducing the risk of bias toward the majority class and enhancing model performance during the training preprocessing step before model training is dimensionality reduction, achieved via Principal Component Analysis (PCA). PCA transforms the dataset into a set of uncorrelated components, effectively reducing the number of input features. This step minimizes the risk of overfitting, enhances computational efficiency, and simplifies the model while retaining essential information from the original features [15]-[16]. Consequently, Principal Component Analysis (PCA) reduces the dataset's dimensionality. Hence, we proposed four scenarios: without SMOTE and PCA, PCA only, SMOTE only, and both for modality-based datasets: image only (unimodal), tabular only (unimodal), and image-tabular (multi-modal). Various machine learning algorithms, including KNN, SVC, Decision Tree, Random Forest, Gradient Boosting, LGBM, Cat Boost, AdaBoost, Bagging Classifier, and MLP Classifier, are then used to train models and detect stunting on the testing data for algorithm comparison. Finally, we evaluate the model using precision, recall, F1-score, and accuracy measurements.

2.4. Modality-based Modelling

Our proposed idea is to model modality-based stunting data involving different modalities to enhance the accuracy and robustness of detection models, as seen in Figure 1. The study includes unimodal and multimodal modeling. The contributions of this research include developing a comprehensive framework for modality-based analysis, utilizing advanced preprocessing techniques, and demonstrating the advantages of integrating multiple modalities for more accurate stunting detection. In the context of our stunting analysis, the proposed scenario explores four different combinations of data balancing and dimension reduction to optimize the model's performance. These include scenarios with and without using SMOTE to balance class distributions and PCA to reduce the complexity of the dataset. Each combination is tested to determine the best approach for handling imbalanced stunting data while maintaining the critical features for analysis. Following this, three modality-based approaches are examined: using only image data (such as health center images to identify visual indicators of stunting) as unimodal analysis, using only tabular data (which includes stunting prevalence statistics from health records) as unimodal analysis, and combining both image and tabular data as multimodal analysis, as multimodal analysis grows in corresponding several different modality data with machine learning modeling. [13], [17]. Tabular-only data only follows the tabular data in the model (age, gender, and weight), while image-only data is obtained from the GLCM result. Image-tabular modeling is achieved by concatenating the output of GLCM from image data and several columns from tabular data, which are age and gender.

2.4.1. Unimodal Modeling

Unimodal modeling refers to the use of a single type of data for modeling. This modeling aims to create a single modality model as traditional data mining and computer vision have done. In this study, we consider two unimodal approaches:

1. Tabular-Only Modeling: This approach uses only tabular data, which includes variables such as age, gender, and weight. The tabular data undergoes preprocessing steps as in tabular analysis in data mining. Various machine-learning algorithms were then applied to train and evaluate the models.
2. Image-Only Modeling: This approach uses only image data. The images are preprocessed through background segmentation and feature extraction with the Gray Level Co-

occurrence Matrix (GLCM). Machine learning algorithms are applied to both model training and evaluation data types.

2.4.2 Multimodal Modeling

Multimodal modeling combines tabular and image data to leverage the strengths of each modality. This approach aims to provide a more comprehensive framework for stunting detection by integrating multiple sources of information. This combines tabular and image data, incorporating their features to improve model performance. The steps involved in multimodal analysis are as follows:

1. Data Preprocessing for each data modality: The tabular data undergoes the same preprocessing steps as in the unimodal analysis. The features used from tabular data are age and gender. We eliminate the weight as it would be predicted through image modality. Then, The image data is preprocessed through background segmentation and feature extraction using GLCM, similar to the image-only analysis.
2. Data Integration: The preprocessed tabular and image data are concatenated to form a combined dataset. This study's tabular data includes age and gender, while the image data provides additional features extracted through GLCM. Integrating these modalities aims to enhance the model's ability to detect stunting by utilizing structured and unstructured data.

2.5. Machine Learning Modelling

The train-test split phase consists of two subsets: training and testing sets. Specifically, 70% of the data is allocated for training the machine learning models, while the remaining 30% is reserved for testing. This joint approach ensures that the models learn from one portion of the data (training set) and are evaluated on a separate, unseen portion (testing set) to gauge their accuracy and generalization to new data. The data split applies to different modalities of image data, tabular data, and a combination of both so that the system can assess how well models perform with each data type. This step is essential to ensure the models can detect stunting by learning from diverse features, including visual data (images) and structured data (tabular). In our study on stunting, we trained several different algorithms to train the datasets, such as KNN, SVM, DT, Ensemble Classifier Method, and Bagging Classifier method. A deeper exploration of these algorithms reveals that each algorithm contributes unique strengths to the modeling process. KNN is a supervised learning method that classifies data based on the closest distance to training data [18].

Support Vector Classification (SVC) is a variant of SVM employed for classification tasks using mathematical formulations. The fundamental concept of SVC can be understood as identifying the optimal hyperplane that separates two distinct classes in the input space. This process involves measuring the margins of the hyperplane and locating the maximum point, which leads to identifying the best-separating hyperplane. The margin refers to the gap between the hyperplane and the closest data points from each class, with these nearest points being known as support vectors [19]. In the context of stunting detection, the decision tree algorithm is another supervised learning method that can classify whether an individual is stunted based on their health attributes. A decision tree model operates by repeatedly dividing the dataset into smaller subsets based on feature values. This process forms a tree-like structure where each internal node signifies a decision based on a particular feature, and each leaf node indicates a classification result (potentially stunted or not stunted) [20]-[21]. For instance, the decision tree algorithm evaluates various health-related features such as height, weight, age, and other factors that might be included in the stunting dataset. It selects the feature that provides the most information about the outcome (i.e., whether a child is stunted) and uses that feature to create a decision rule at the root node.

Ensemble techniques are particularly effective in improving the accuracy and reliability of predictions. [22]. RF uses multiple decision trees to enhance generalization and reduce overfitting. [23]. GB, on the other hand, builds trees sequentially, with each tree correcting errors from the previous one [20], [22]. While GB is computationally more intensive than simpler algorithms like decision trees, it often delivers superior predictive accuracy, especially in complex datasets with non-linear relationships and noisy features. [23]. LGBM, a recent innovation, optimizes GB by reducing training time and improving efficiency through leaf-wise growth and histogram-based techniques [26]. CB is designed to handle categorical variables effectively, making it particularly useful in real-world data scenarios with abundant categorical features. [27], [28]. Finally, AB focuses on adjusting model weights based on misclassified instances, boosting weak learners iteratively to improve performance. [22], [29].

Bagging (short for Bootstrap Aggregating) is a widely recognized family of ensemble learning techniques, particularly effective in decision tree classification tasks. This method leverages the manipulation of training samples to improve model performance by combining multiple predictors. Bagging operates as a regular ensemble method, where several individual models are built independently of one another and then aggregated using methods like weighted averaging or majority voting to generate a final prediction. The primary goal of Bagging is to reduce model variance by averaging the predictions of diverse models, which is especially useful for high-variance algorithms like decision trees. [30].

MLP is a feedforward neural network consisting of an input layer, one or more hidden layers, and an output layer. MLP neural networks can be especially beneficial when applied to multimodal data (i.e., the combination of image and tabular data), as they could learn hierarchical feature representations. For example, in image data, MLP can learn abstract features that might indicate stunting from a child's physical appearance while simultaneously learning from tabular data such as height, weight, and age. The ability of MLP to generalize well on unseen data could potentially improve classification accuracy, especially when combined with techniques like SMOTE for data balancing and PCA for dimensionality reduction. Moreover, MLP can be further optimized with techniques like hyperparameter tuning (adjusting the number of layers, nodes per layer, learning rate, etc.), dropout to prevent overfitting, and batch normalization for stabilizing and accelerating training. [31].

3. RESULT AND DISCUSSION

This study aims to create a novel framework for detecting the potential of stunting using modality-based modeling, namely unimodal and multimodal, with data balancing and dimensionality reduction. Unimodal analysis used tabular-only (only tabular dataset) and image-only (only images dataset). As recommended in the KIA book, the tabular-only analysis uses several variables utilized in Indonesia to control children's nutrition, such as age, gender, and weight. The image-only analysis uses one child's image as a predictor. Multimodal analysis used image-tabular (image and tabular analysis). In image-tabular, we only use age and gender in tabular form paired with the image as a substitution for weight in the tabular-only analysis. GLCM extracts image features with regular parameters: Contrast, Correlation, Homogeneity, Energy, Entropy, and Angular Second Moment (ASM). These are included in the angles 0, 45, 90, and 135 degrees. Thus, there are 24 extracted features of the image. SMOTE is used for data balancing (oversampling) to handle imbalanced data. Then, PCA is utilized to analyze the effectiveness of dimensionality reduction for further data analysis. KNN, SVM, DT, RF, GB, LGBM, CB, AB, BC, and MLP do machine learning modeling. The default hyperparameter, defined in the SKLearn package modeling on Python Programming, is the default hyperparameter. As the proposed scenario mentioned in the previous section, these scenarios run with three modalities: tabular-only, image-only, and tabular-image analysis. F1-Score from stunted children is computed to maintain the control of stunted children predicted and the majority of the dataset, which non-stunted children own.

3.1. Without SMOTE and PCA Analysis

The first scenario in this experiment involves running the three modalities without applying data balancing (SMOTE) and dimensionality reduction. Figure 2 shows the performance of the machine learning models using F1-Score metrics: Figure 2(a) for tabular-only, Figure 2(b) for image-only, and Figure 2(c) for image-tabular modeling. The results indicate that tabular-only modeling achieves the best F1-Score, over 50%, with ensemble models being the top performers. Image-only and tabular-only modeling have similar results at 32% due to the need for more prediction using all features as predictors. Tree-based models, such as RF, GB, and LGBM, dominate as the best models in image-only and tabular-only scenarios, while traditional models like KNN, SVM, and MLP underperform. Table 1 summarizes the best models across the modalities, showing that accuracy exceeds 80% in the tabular-only model despite lower precision and recall. The effectiveness of specific algorithms is influenced by the quality and volume of data and the importance of features in the dataset.

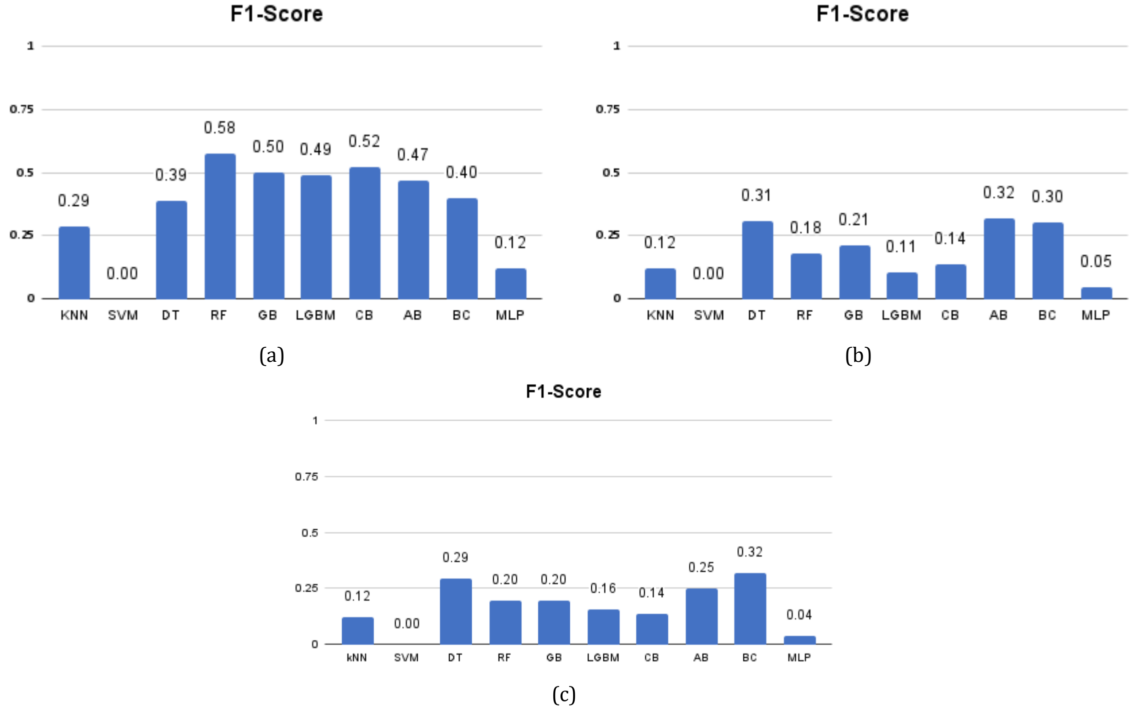


Figure 2. F1-Score Comparison in Scenario 1 between modality: (a) Tabular-Only, (b) Image-Only, (c) Image-Tabular

Table 1. Summary of The Best Metrics in Scenario 1

Modality	**Best Model**	**F1-Score**	**Accuracy**	**Precision**	**Recall**
Tabular-only	RF	0.58	0.82	0.61	0.54
Image-only	AB	0.32	0.78	0.32	0.32
Image-Tabular	BC	0.32	0.79	0.39	0.27

3.2. Without SMOTE and With PCA Analysis

The second scenario examines the impact of dimensionality reduction using PCA on modality analysis without data balancing. Figure 3 presents the F1-Score results: Figure 3(a) for Tabular-Only, Figure 3(b) for Image-Only, and Figure 3(c) for Image-Tabular modeling. PCA does not significantly affect Image-Only and Image-Tabular models, as reducing features decreases the model's ability to detect stunting correctly, with the majority class dominating incorrect predictions. However, PCA positively impacts the Tabular-Only model, improving recall to 73% even with reduced features. Table 2 summarizes the best metrics, showing that Tabular-Only modeling achieves the best results. Other modalities still exhibit lower recall and precision, indicating higher false negatives. Ensemble models like LGBM and AB perform best in the tabular models, with PCA enhancing their ability to identify significant features.

p-ISSN: 2528-1682
e-ISSN: 2527-9165

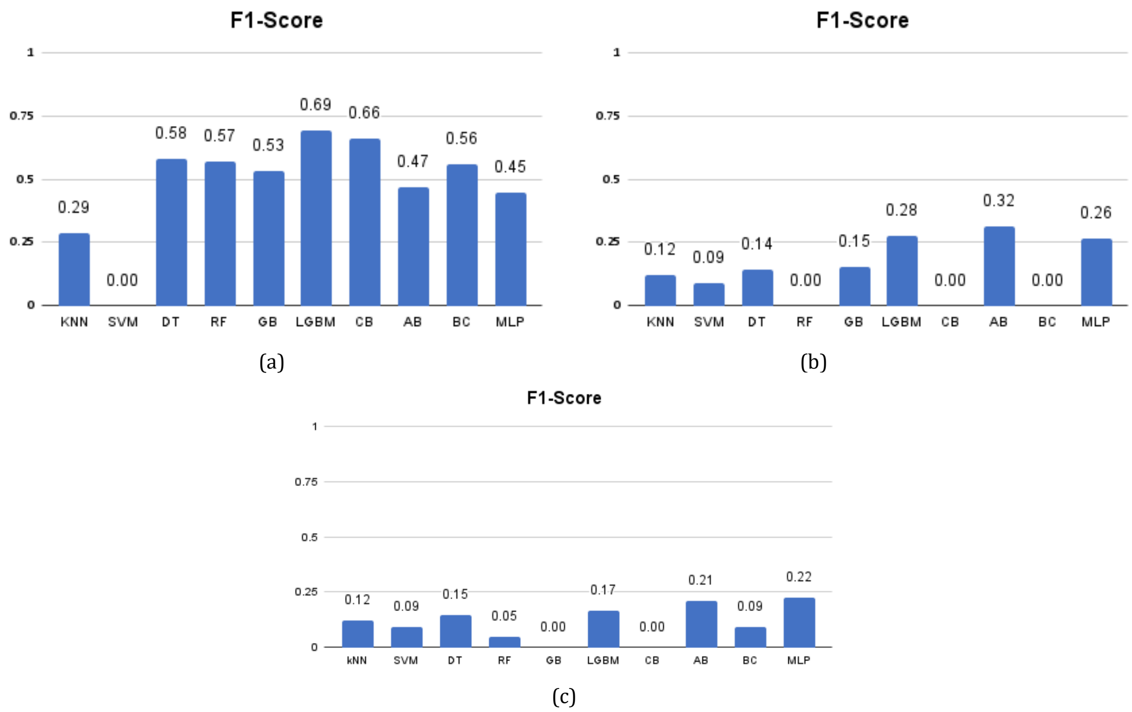


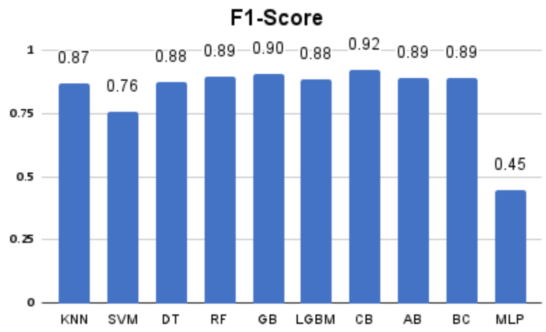
Figure 3. F1-Score Comparison in Scenario 2 between modality: (a) Tabular-Only, (b) Image-Only, (c) Image-Tabular

Table 2. Summary of The Best Metrics in Scenario 2

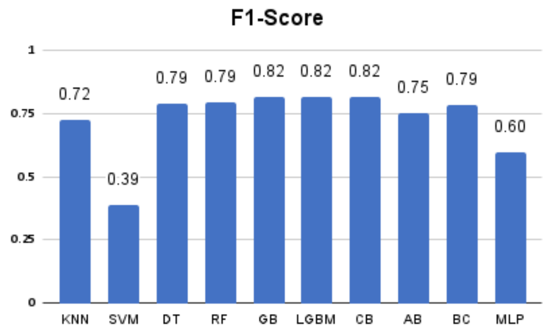
Modality	Best Model	F1-Score	Accuracy	Precision	Recall
Tabular-only	LGBM	0.69	0.89	0.65	0.73
Image-only	AB	0.31	0.73	0.30	0.32
Image-Tabular	MLP	0.22	0.74	0.21	0.23

3.3. With SMOTE and Without PCA Analysis

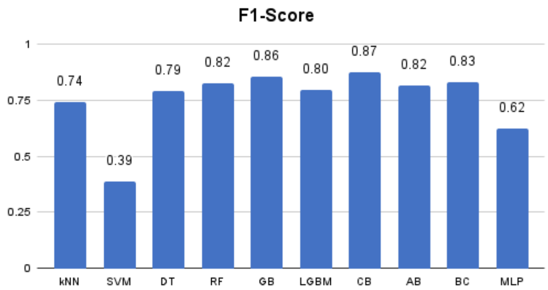
In the following experiment, training trials were conducted using SMOTE (Synthetic Minority Over-sampling Technique) to handle imbalanced data without applying PCA for dimensionality reduction. The Tabular-Only dataset outperformed both the Image-Only and Image-Tabular datasets, achieving the highest average F1-Score of 0.83. However, as shown in Figure 4(b), the Image-Only dataset exhibited more stable and consistent results with a lower standard deviation of 0.13, indicating more reliable model outcomes. CatBoost displayed the highest resilience across diverse dataset structures, with an average F1-Score of 0.87, suggesting its effectiveness in achieving robust classification performance. For the Tabular-only dataset, CB attained an accuracy of 90% and an F1-Score of 92%, while Gradient Boosting was most suitable for the Image-only dataset, achieving an accuracy and F1-Score of 82%. In the combined Image-Tabular dataset, CatBoost led with an accuracy of 86% and an F1-Score of 87%, highlighting its adaptability in integrating multiple data modalities. These results, summarized in Table 3, suggest that CB is particularly suited for handling complex and varied dataset characteristics, especially within the context of the Stunting dataset using SMOTE-based feature engineering.



(a)



(b)



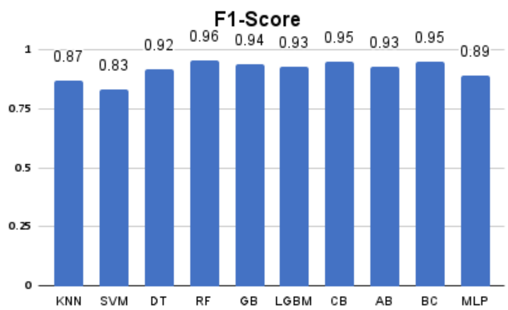
(c)

Figure 4. F1-Score Comparison in Scenario 3 between modality: (a) Tabular-Only, (b) Image-Only, (c) Image-Tabular

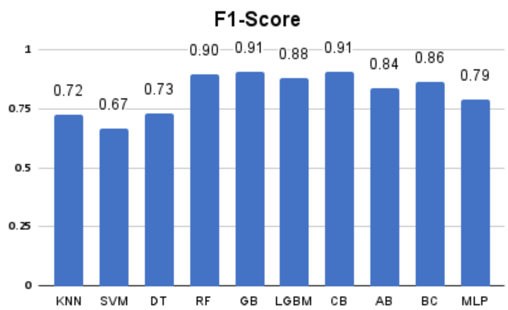
Table 3. Summary of The Best Metrics in Scenario 3

Modality	Best Model	F1-Score	Accuracy	Precision	Recall
Tabular-only	CB	0.92	0.90	0.92	0.92
Image-only	GB	0.82	0.82	0.77	0.88
Image-Tabular	CB	0.87	0.86	0.83	0.91

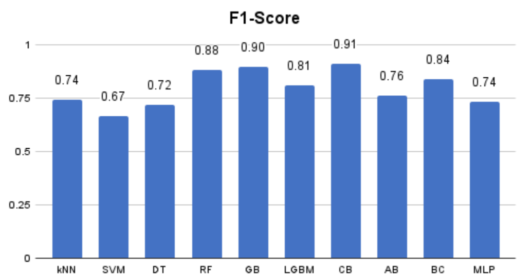
3.4. With SMOTE and PCA Analysis



(a)



(b)



(c)

Figure 5. F1-Score Comparison in Scenario 4 between modality: (a) Tabular-Only, (b) Image-Only, (c) Image-Tabular

In the final experimental scenario, SMOTE feature engineering was combined with Principal Component Analysis (PCA) for dimensionality reduction. The Tabular-only dataset demonstrated consistently balanced model performance, achieving an average F1-Score of 0.92, attributed to the simplicity of its feature structure. Figure 5(a) shows a low standard deviation of 0.039 for the Tabular-only dataset, indicating greater consistency in model outcomes. Figures 5(b) and 5(c) provide another result from image-only and image-tabular modeling. CB emerged as the most robust method, achieving an average F1-Score of 0.923, followed closely by Random Forest with an F1-Score of 0.913. The trial results using SMOTE and PCA demonstrated superior model performance compared to trials without these techniques, as shown in Table 4. Specifically, Random Forest achieved the highest accuracy on the Tabular-only (94%) and Image-Tabular (88%) modeling, while CatBoost led in the Image-only dataset with 91% accuracy. These findings suggest that ensemble methods like Random Forest and CatBoost are well-suited for stunting datasets, and the combined application of SMOTE and PCA enhances model robustness and accuracy across varying dataset structures.

Table 4. Summary of The Best Metrics in Scenario 3

Modality	Best Model	F1-Score	Accuracy	Precision	Recall
Tabular-only	RF	0.96	0.94	0.94	0.98
Image-only	CB	0.91	0.91	0.88	0.94
Image-Tabular	RF	0.90	0.88	0.92	0.88

3.5. *Discussion*

This study evaluates the effectiveness of unimodal and multimodal frameworks in detecting stunting potential through the lens of advanced data preprocessing techniques and machine learning algorithms. The discussion focuses on the implications of using data balancing, dimensionality reduction, and different modalities for model performance. It provides insights into the broader applicability of these methods in similar contexts. The results underscore the superiority of the Tabular-Only dataset in achieving high predictive performance. When SMOTE was applied without PCA, the Tabular-Only modality achieved the highest F1-Score of 0.83 and demonstrated consistent performance across classification models. This suggests that tabular features, particularly age and gender, are robust predictors of stunting when used in isolation. The integration of ensemble models, such as CatBoost and Random Forest, further enhanced the classification outcomes, reflecting their capability to leverage structured tabular data effectively.

Conversely, the Image-Only modality exhibited lower predictive accuracy but demonstrated more excellent stability, as evidenced by its low standard deviation of 0.13. This suggests that while image features alone may lack the richness needed for high accuracy in stunting prediction, they contribute to a more consistent model output. This stability can be advantageous in scenarios prioritizing reliability over maximum accuracy. The combined Image-Tabular modality showed moderate performance, indicating the potential benefits and challenges of integrating heterogeneous data types. Although CatBoost achieved a high F1-Score of 0.87 in this modality, the overall performance lagged behind the Tabular-Only dataset. This outcome highlights the complexity of multimodal integration, where balancing the contributions of disparate data types remains a significant challenge.

The use of PCA for dimensionality reduction yielded mixed results. PCA improved recall to 73% for the Tabular-Only dataset, correctly identifying stunted children with reduced features. This demonstrates PCA's utility in identifying and preserving critical features while reducing data complexity. However, PCA's impact on Image-Only and Image-Tabular datasets was less favorable, as the reduced feature set negatively affected the model's ability to detect stunting accurately. This suggests that dimensionality reduction techniques must be carefully tailored to the specific characteristics of the data modality. The findings emphasize the importance of modality-specific preprocessing methods. For instance, SMOTE effectively mitigates class imbalance in all modalities, but its combination with PCA significantly enhances performance only in structured tabular datasets. This indicates that the interplay between data balancing and dimensionality reduction is context-dependent and must be optimized for each dataset type. Furthermore, the robustness of ensemble models, particularly CatBoost, across all modalities highlights their versatility in handling diverse data structures. CatBoost's high performance with the Tabular-Only dataset (F1-Score of 0.92) and its adaptability to multimodal data underline its potential as a go-to algorithm for complex classification tasks. A notable insight is the consistency observed in the Image-Only dataset. While this modality underperformed in absolute terms, its stability suggests that image features may be complementary predictors when integrated with other data types. Future studies could explore advanced feature extraction techniques, such as deep learning-based methods, to enhance the predictive power of image features. While the study provides compelling evidence for the efficacy of the proposed framework, certain limitations should be acknowledged. The reliance on default hyperparameters in machine learning models may have constrained their performance potential. Future research could investigate the impact of hyperparameter tuning and the incorporation of advanced modeling techniques, such as neural networks, to further enhance predictive accuracy. Additionally, the study's focus on a specific dataset structure limits the generalizability of its findings. Expanding the framework to include diverse datasets with varying feature distributions and data types could validate its robustness and scalability. Moreover, integrating domain-specific knowledge into feature selection processes could further refine model performance.

4. CONCLUSION

This study addresses the critical issue of early stunting detection by developing a comprehensive modality-based data mining framework. The high prevalence of stunting in Indonesia, significantly above the WHO standard, underscores the urgent need for effective detection methods. Traditional approaches focusing on unimodal data have limitations in comprehensiveness and accuracy. Our research addresses this problem by integrating image and tabular data, leveraging advanced data preprocessing techniques such as SMOTE for data balancing and PCA for dimensionality reduction. The main contributions of this research include the development of a novel modality-based framework encompassing both unimodal (tabular-only and image-only) and multimodal (image-tabular) modeling, the application of SMOTE and PCA to address class imbalance and high dimensionality, and the comparison of various machine learning algorithms to identify the best models for each modality. The first scenario, which did not apply SMOTE or PCA, resulted in lower F1 scores: 0.58 for tabular data, 0.32 for image data, and 0.32 for image-tabular data. When PCA was applied without SMOTE, there was an improvement in the tabular data's F1-score to 0.69, though performance for the other modalities remained low. Significant improvements were seen when SMOTE was applied without PCA, with F1-scores rising to 0.92 for tabular data, 0.82 for image data, and 0.87 for image-tabular data. The best results were achieved in the final scenario, where SMOTE and PCA were utilized, yielding F1 scores of 0.96 for tabular data, 0.91 for image data, and 0.90 for image-tabular data. These results underscore the importance of combining data balancing and dimensionality reduction techniques to improve the accuracy of stunting detection, especially in tabular and multimodal data setups. Future research could explore additional sampling methods to address class imbalance, such as random over-sampling and under-sampling. Additionally, incorporating more data types, such as text data from medical records or sensor data, could extend the multimodal analysis framework, creating a more comprehensive approach to stunting detection. By addressing these areas, future research can build on the findings of this study to develop even more effective and robust stunting detection models.

ACKNOWLEDGEMENTS

This research was funded by Telkom University under the Skema Penelitian Dasar Tahun Anggaran 2024 No. 301/LIT06/PPM-LIT/2024. We express our gratitude to our colleagues at Telkom University's Surabaya Campus for their valuable insights and expertise, which significantly contributed to the research. We also thank Puskesmas Gedangan in Sidoarjo, Indonesia, for sharing their knowledge on stunting potential and providing a site for data collection.

REFERENCES

[1] W. Hadikurniawati, K. Dwi Hartomo, I. Sembiring, H. Dwi Purnomo, and A. Iriani, "Triangular fuzzy numbers-based MADM for selecting pregnant mothers at risk of stunting," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 3, pp. 579–585, Jun. 2023, doi: 10.29207/resti.v7i3.4966.

[2] L. A. Jokhu, A. Syauqy, L. Y. Lin, F. F. Dieny, and A. Rahadiyanti, "Determinants of stunting among children 6–23 months: a population-based study in Indonesia," *Nutr Food Sci*, 2024, doi: 10.1108/NFS-01-2024-0025.

[3] I. K. Hasan, Nurwan, Nur Falaq, and Muhammad Rezky Friesta Payu, "Optimization Fuzzy Geographically Weighted Clustering with Gravitational Search Algorithm for Factors Analysis Associated with Stunting," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 1, pp. 120–128, Feb. 2023, doi: 10.29207/resti.v7i1.4508.

[4] Y. Yuliana, P. Paradise, and M. Qulub, "Detection Of Children's Nutritional Status Using Machine Learning With Logistic Regression Algorithm," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 10, no. 2, pp. 267–274, Mar. 2024, doi: 10.33330/jurteksi.v10i2.2973.

[5] R. Gustriansyah, N. Suhandi, S. Puspasari, and A. Sanmorino, "Machine Learning Method to Predict the Toddlers' Nutritional Status," *JURNAL INFOTEL*, vol. 16, no. 1, Jan. 2024, doi: 10.20895/infotel.v15i4.988.

[6] A. A. Ningrum and Y. Ikawati, "Early Detection of Stunting Based on Feature Engineering and Machine Learning Algorithm Approach," *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 6, no. 3, pp. 147–155, 2024, doi: 10.35882/ijeeemi.v6i3.6.

[7] A. A. Permana, B. Raharja, and A. T. Perdana, "Artificial Intelligence for Diagnosing Child Stunting: A Systematic Literature Review," *Journal of System and Management Sciences*, vol. 13, no. 6, pp. 605–621, 2023, doi: 10.33168/JSMS.2023.0635.

[8] R. Saragih, E. N. Saputra, D. Setiadikarunia, and J. J. Jarden, "Perbandingan Teknik Klasifikasi Fast Null-space Based Linear Discriminant Analysis (FNLDA) dan Direct Linear Discriminant Analysis (DLDA) dalam Pengenalan Citra Multimodal Wajah atau Pembuluh Darah di Telapak Tangan," *Jurnal Ecotipe (Electronic, Control, Telecommunication, Information, and Power Engineering)*, vol. 9, no. 1, pp. 40–48, Apr. 2022, doi: 10.33019/jurnalecotipe.v9i1.2867.

[9] Y. Wang *et al.*, "Geometric Correspondence-Based Multimodal Learning for Ophthalmic Image Analysis," *IEEE Trans Med Imaging*, vol. 43, no. 5, pp. 1945–1957, May 2024, doi: 10.1109/TMI.2024.3352602.

[10] K. N. Singh, O. P. Singh, A. K. Singh, and A. K. Agrawal, "WatMIF: Multimodal Medical Image Fusion-Based Watermarking for Telehealth Applications," *Cognit Comput*, vol. 16, no. 4, pp. 1947–1963, Jul. 2024, doi: 10.1007/s12559-022-10040-4.

[11] C. Cui *et al.*, "Deep multimodal fusion of image and non-image data in disease diagnosis and prognosis: a review," *Progress in Biomedical Engineering*, vol. 5, no. 2, Apr. 2023, doi: 10.1088/2516-1091/acc2fe.

[12] F. Behrad and M. Saniee Abadeh, "An overview of deep learning methods for multimodal medical data mining," Aug. 15, 2022, *Elsevier Ltd.* doi: 10.1016/j.eswa.2022.117006.

[13] C. Liu, Z. Mao, T. Zhang, A. A. Liu, B. Wang, and Y. Zhang, "Focus Your Attention: A Focal Attention for Multimodal Learning," *IEEE Trans Multimedia*, vol. 24, pp. 103–115, 2022, doi: 10.1109/TMM.2020.3046855.

[14] D. Elreedy, A. F. Atiya, and F. Kamalov, "A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning," *Mach Learn*, vol. 113, no. 7, pp. 4903–4923, Jul. 2024, doi: 10.1007/s10994-022-06296-4.

[15] M. Greenacre *et al.*, "Economics Working Paper Series Principal component analysis Principal Component Analysis," Barcelona, Jan. 2023.

[16] U. N. Wisesty, T. A. B. Wirayuda, F. Sthevanie, and R. Rismala, "Analysis of Data and Feature Processing on Stroke Prediction using Wide Range Machine Learning Model," *Jurnal Online Informatika*, vol. 9, no. 1, pp. 29–40, Apr. 2024, doi: 10.15575/join.v9i1.1249.

[17] Y. Yin, H. Wang, S. Liu, J. Sun, P. Jing, and Y. Liu, "Internet of Things for Diagnosis of Alzheimer's Disease: A Multimodal Machine Learning Approach Based on Eye Movement Features," *IEEE Internet Things J*, vol. 10, no. 13, pp. 11476–11485, Jul. 2023, doi: 10.1109/JIOT.2023.3245067.

[18] M. Ula, S. Fachrurrazi, and R. Achmad Rizal, "Implementation Of Data Mining Models With Algorithms K-Nearest Neighbor In Monitoring The Nutritional Status Of Children And Stunting," *Journal of Information Systems and Computer Science Prima)*, vol. 6, no. 2, 2023.

[19] M. I. Irawan, M. H. Z. Al Faroby, and A. P. Dyah Nurhayati, "In Silico Analysis Using Hybrid Support Vector Machine and Second Order of Markov Chain for Multiple Sequence Alignment to Identify the Types of Leukaemia," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Nov. 2019. doi: 10.1088/1742-6596/1366/1/012052.

[20] A. Nugroho, H. L. Hendric, S. Warnars, F. L. Gaol, and T. Matsuo, "Trend of Stunting Weight for Infants and Toddlers Using Decision Tree," *Int J Appl Math (Sofia)*, vol. 52, no. 1, Mar. 2022.

[21] Y. Setiawan, O. A. Permata, and M. P. Yuda, "Decision Tree based Data Modelling for First Detection of Thalassemia Major," *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 13, no. 1, pp. 49–56, Feb. 2024, doi: 10.32736/sisfokom.v13i1.1949.

[22] I. D. Mienye and Y. Sun, "A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects," 2022, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2022.3207287.

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

[24] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Front Neurorobot*, vol. 7, no. 21, 2013, doi: 10.3389/fnbot.2013.00021.

[25] M. H. Z. AlFaroby, H. N. Fadhilah, S. Amiroch, and R. S. Hidayat, "XGB-Hybrid Fingerprint Classification Model for Virtual Screening of Meningitis Drug Compounds Candidate," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, Jun. 2022, doi: 10.22219/kinetik.v7i2.1424.
[26] D. A. McCarty, H. W. Kim, and H. K. Lee, "Evaluation of light gradient boosted machine learning technique in large scale land use and land cover classification," *Environments - MDPI*, vol. 7, no. 10, pp. 1–22, Oct. 2020, doi: 10.3390/environments7100084.
[27] D. Niu, L. Diao, Z. Zang, H. Che, T. Zhang, and X. Chen, "A Machine-Learning Approach Combining Wavelet Packet Denoising with Catboost for Weather Forecasting," *Atmosphere (Basel)*, vol. 12, no. 12, Dec. 2021, doi: 10.3390/atmos12121618.
[28] J. Tanha, Y. Abdi, N. Samadi, N. Razzaghi, and M. Asadpour, "Boosting methods for multi-class imbalanced data classification: an experimental review," *J Big Data*, vol. 7, no. 1, Dec. 2020, doi: 10.1186/s40537-020-00349-y.
[29] F. Wang, Z. Li, F. He, R. Wang, W. Yu, and F. Nie, "Feature Learning Viewpoint of Adaboost and a New Algorithm," *IEEE Access*, vol. 7, pp. 149890–149899, 2019, doi: 10.1109/ACCESS.2019.2947359.
[30] H. Jafarzadeh, M. Mahdianpari, E. Gill, F. Mohammadimanesh, and S. Homayouni, "Bagging and boosting ensemble classifiers for classification of multispectral, hyperspectral and polSAR data: A comparative evaluation," *Remote Sens (Basel)*, vol. 13, no. 21, Nov. 2021, doi: 10.3390/rs13214405.
[31] F. M. Amin, D. Candra, and R. Novitasari, "Identification of Stunting Disease using Anthropometry Data and Long Short-Term Memory (LSTM) Model," *Computer Engineering and Applications*, vol. 11, no. 1, 2022.