JURNAL INFOTEL
VOL. 16, NO. 1, FEBRUARY 2024, PP. 32–43. DOI:10.20895/INFOTEL.V15I4.988

RESEARCH ARTICLE

# Machine Learning Method to Predict the Toddlers' Nutritional Status

Rendra Gustriansyah*, Nazori Suhandi, Shinta Puspasari, and Ahmad Sanmorino

Faculty of Computer Science, Universitas Indo Global Mandiri, Palembang 30129, Indonesia

*Corresponding email: rendra@uigm.ac.id

*Received: July 6, 2023; Revised: September 22, 2023; Accepted: October 5, 2023.*

**Abstract:** Malnutrition is one of the leading health problems experienced by toddlers in various countries. Based on the 2022 Indonesian Nutritional Status Survey results, malnutrition in children under five in Indonesia is higher than the average malnutrition in Africa and globally. Therefore, a way is needed to predict the nutritional status of children under five early and quickly so that the Government (through the District Health Office) can immediately provide the necessary treatment. This study aims to predict or classify the toddlers' nutritional status based on age, body mass index (BMI), weight, and body length using various machine learning (ML) methods, namely naïve Bayes, linear discriminant analysis, decision tree, k-nearest neighbor, random forest, and support vector machine. The predictive performance of each ML method was evaluated based on accuracy, sensitivity, specificity, the area under the curve (AUC), and Cohen's kappa coefficient (CKC). The test results show that the random forest method is the most recommended for predicting toddlers' nutritional status with accuracy, sensitivity, specificity, AUC, and CKC values: 0.9737, 0.9500, 0.9881, 0.9990, and 0.9609, respectively. The study's contribution is to obtain information about toddlers' nutritional status more easily.

**Keywords:** classification, machine learning, nutritional status, prediction, toddlers

## 1 Introduction

Malnutrition is one of the major health problems experienced by children under five in many countries. Malnutrition can affect the overall growth and development of children under five. Based on the 2022 Indonesian Nutrition Status Survey (SSGI) results, around 17.1 % of children under five in Indonesia were underweight, and 7.7 % were malnourished (wasting). This undernutrition rate increased by about 0.1 % compared to 2021. Meanwhile, malnutrition increased by 0.6 % [1]. Malnutrition in toddlers in Indonesia is higher than the average malnutrition in Africa (5.8 %) and globally (6.8 %) [2]. This situation is made worse by the lack of public awareness of the nutritional value, the difficulty in obtaining wholesome food, the existence of cases of poverty, and the scarcity of food.

If not appropriately addressed, malnutrition in children under five can potentially cause serious health problems, increasing the risk of developing various long-term health problems, such as diabetes, high blood pressure, stroke, behavioral disorders, and even death. Therefore, it is necessary to detect the nutritional status of children under five years of age early and quickly so that the government (through the District Health Office) can immediately provide the necessary treatment. The machine learning (ML) method can rapidly predict the toddlers' nutritional status by classifying them based on age, BMI, weight, and length. It will impact the ease of obtaining information about the toddlers' nutritional status.

Several ML methods are used to classify/predict malnutrition or nutritional status in toddlers, including the naïve Bayes (NB) method [3–6], logistic regression [7], k-nearest neighbor (kNN) [4, 5, 8], decision tree (DT) [6, 9–11], support vector machine (SVM) [12], and learning vector quantization [13]. Several studies compare several ML methods to classify malnutrition in toddlers [4, 14–19]. However, no research in Indonesia uses several ML methods to predict several nutritional statuses (multi-class) in toddlers, such as malnutrition, undernutrition, normal, and overnutrition. Therefore, this study aims to predict the toddlers' nutritional status using various ML methods. The ML methods considered include NB, linear discriminant analysis (LDA), DT, kNN, random forest (RF), and SVM. These methods are a combination of simple linear (NB, LDA), nonlinear (DT, k-NN), and complex nonlinear (RF, SVM) methods. Furthermore, the predictive performance of these methods is evaluated based on their accuracy, sensitivity, specificity, the area under curve (AUC), and Cohen's kappa coefficient (CKC). The classification method with the best evaluation results can be recommended as an ML method to predict the nutritional status of children under five so that the Government can immediately provide the necessary treatment.

## 2 Research Methods

This section discusses the research stages, cata collection, data preprocessing and splitting, parameter tuning, classification/prediction methods, and evaluation.

### 2.1 Research Stages

Figure 1 shows the stages of this research, and the R programming tool was used starting from the data preprocessing to the results evaluation stage.

### 2.2 Data Collection

The dataset used in this study was obtained from the Kaggle site [20]. The dataset consists of 200 toddler data with seven variables: name, gender, age, weight, height, BMI, and toddler status. However, two variables, name and gender, will not be included in the classification process because they are not involved in the calculation. So, only five variables from the dataset are used for training and testing. The target variable (toddler status) is a

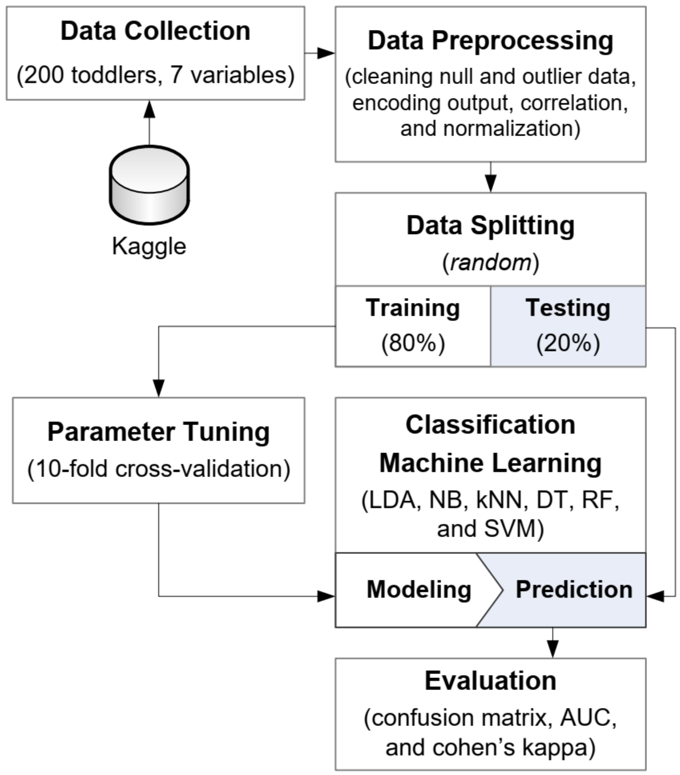


Figure 1: Research stages.

multi-class variable that contains malnutrition, undernutrition, overnutrition, and normal classes.

## 2.3 Data Preprocessing and Splitting

Before the dataset is randomly split into 80 % data for training and the remaining 20 % for testing, it is crucial to clean the dataset from 'null' data, outliers, and duplicates. Correlation between variables will be measured using the Spearman rank correlation coefficient. Next, the output variable (under-five status) is encoded as category (factor) type, and all input variables are normalized/standardized by scaling and centering. The scaling (z-score) and centering (data reduction to the mean) results are directly applied to the training process of each ML classification method.

## 2.4 Parameter Tuning

Parameter tuning aims to organize model training systematically by dividing the dataset into particular and different sizes to produce optimal prediction accuracy. The dividing sample data in this study uses the stratified k-fold cross-validation (CV) method with a value of k = 10. This method will perform resampling for all possible cases from the dataset to minimize the estimation bias level. Therefore, this method is widely used in various classification studies [21].

## 2.5 Classification / Prediction Methods

After examining several previous works, six ML methods that are a combination of simple linear (LDA, NB), nonlinear (k-NN, DT), and complex nonlinear (RF, SVM) methods are considered in this work. LDA is a statistical method used to analyze discriminants in data. The goal is to find a linear combination of features that can best discriminate between the various classes in the given data. To maximize the difference between classes, LDA projects the data to a lower dimensional feature space [22].

NB is a classifier that uses probability and statistical methods to predict future opportunities based on experience [23]. In other words, NB is a classification method based on Bayes' theorem with the assumption that the features used in classification are conditionally independent. The NB method is computationally fast and efficient and has high accuracy despite the small training data amount [24].

One of the simple yet effective classification methods is kNN. The basic principle of kNN is that a sample will be classified based on the majority of its nearest neighbor categories. The "k" in k-NN indicates the number of nearest neighbors to be used in the classification process. The kNN method only stores training data as a knowledge base, so it does not require a complex learning process. The kNN method can work optimally for data with noise [5].

The most common and easily understood classification and regression method is the DT. A tree structure is used by DT to represent decisions and their consequences, where each node in the tree represents a decision to be made based on observed features. DT methods are easy to interpret and can provide a good visual understanding of the train of thought that the model uses for classification (categorical) or regression (numerical) [25].

RF is a group learning method that combines multiple DTs into a more stable and accurate model. DTs are built randomly, and each tree provides a classification for class labeling. The RF method can generate predictions based on a combination of individual predictions given by each tree in the group. Besides being able to handle large and complex datasets, RF can also handle various data, such as numerical and categorical data.

SVM is a kernel-based supervised ML method widely used for regression and classification. By maximizing the margin between classes, the SVM method creates a hyperplane that best divides the training observations based on their class labels. SVM can handle high-dimensional data and nonlinear feature spaces. SVM can also produce models with strong overfitting and good predictability [26].

## 2.6 Evaluation

Evaluation of the prediction results aims to measure the performance of each ML classification method. The performance of each ML method in this study will be evaluated/measured using the confusion matrix (CM) [27], the area under the curve [28], and CKC [27,28].

### 2.6.1 Confusion matrix

According to the CM, there are four possible prediction results, namely true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Meanwhile, the metrics that will be used to evaluate the performance of each ML method in this study are [29]:

$$\text{Accuracy} = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{1}$$

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \tag{2}$$

$$\text{Specifity} = \frac{TN}{(TN + FP)} \tag{3}$$

The accuracy metric describes how accurately the model can classify correctly. Accuracy is valuable as it shows how closely the predicted value matches the true value. The accuracy value can be obtained with (1). In this study, the accuracy value indicates the proportion of all under-fives who were correctly predicted to be malnourished, undernourished, overnourished, or normal.

The ratio of the number of correctly predicted positive data to all positively labeled data is known as the sensitivity metric. The sensitivity value is calculated using (2). The percentage of children under five who are correctly predicted to be malnourished, undernourished, overnourished, or normal out of all children under five who are malnourished/undernourished/overnourished/normal is represented by the sensitivity value.

The specificity metric measures the percentage of negative correct predictions compared to all negatively labeled data. The specificity value shows the toddler percentage who are correctly predicted not to be malnourished, undernourished, overnourished, or normal out of all toddlers who are not malnourished/undernourished/overnourished/ normal. Equation (3) can be used to calculate the specificity value.

### 2.6.2 The area under curve

It is an evaluation metric to measure the overall performance of multi-class classification methods with threshold value settings. AUC measures the ability of a model to distinguish between two classes or groups. In this multi-class study, one class is considered as the "positive" class, while the remaining class is considered as the "negative" class. The higher the AUC value (closer to 1), the more accurate the model is in predicting the class. The discrimination level of AUC can be measured using the data in Table 2 [28].

Table 1: The AUC discrimination level

| AUC value | Discrimination Level |
|---|---|
| < 0.50 | No discrimination |
| 0.51 – 0.70 | Poor |
| 0.71 – 0.80 | Acceptable |
| 0.81 – 0.90 | Excellent |
| 0.91 – 1.00 | Outstanding |

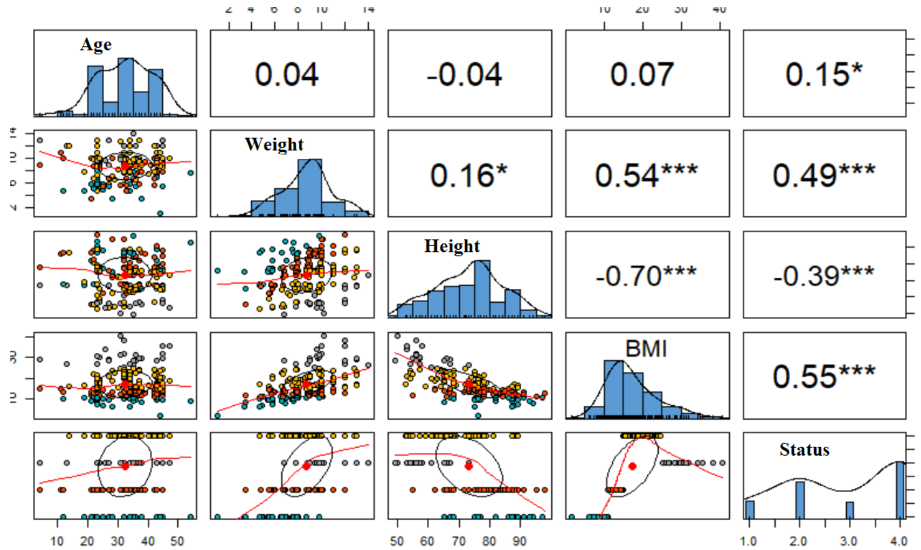


Figure 2: Exploratory data analysis.

### 2.6.3 Cohen's Kappa coefficient

It measures the inter-rater reliability (agreement between annotators) on classification problems. The CKC can be applied to both multi-class and unbalanced classifications. With coefficient values ranging from 0 (poor agreement) to 1 (perfect agreement), as listed in Table 2, the CKC is only applied to qualitative (categorical) data measurement results [27].

Table 2: Agreement level of the CKC [28]

| CKC Value | Agreement Level |
|---|---|
| < 0 | Poor |
| 0.01 – 0.20 | Slight |
| 0.21 – 0.40 | Fair |
| 0.41 – 0.60 | Moderate |
| 0.61 – 0.70 | Substantial |
| 0.81 – 0.99 | Almost perfect |

## 3 Result

This section discusses dataset analysis results, confusion matrix with k-Fold CV, AUC, and CKC.

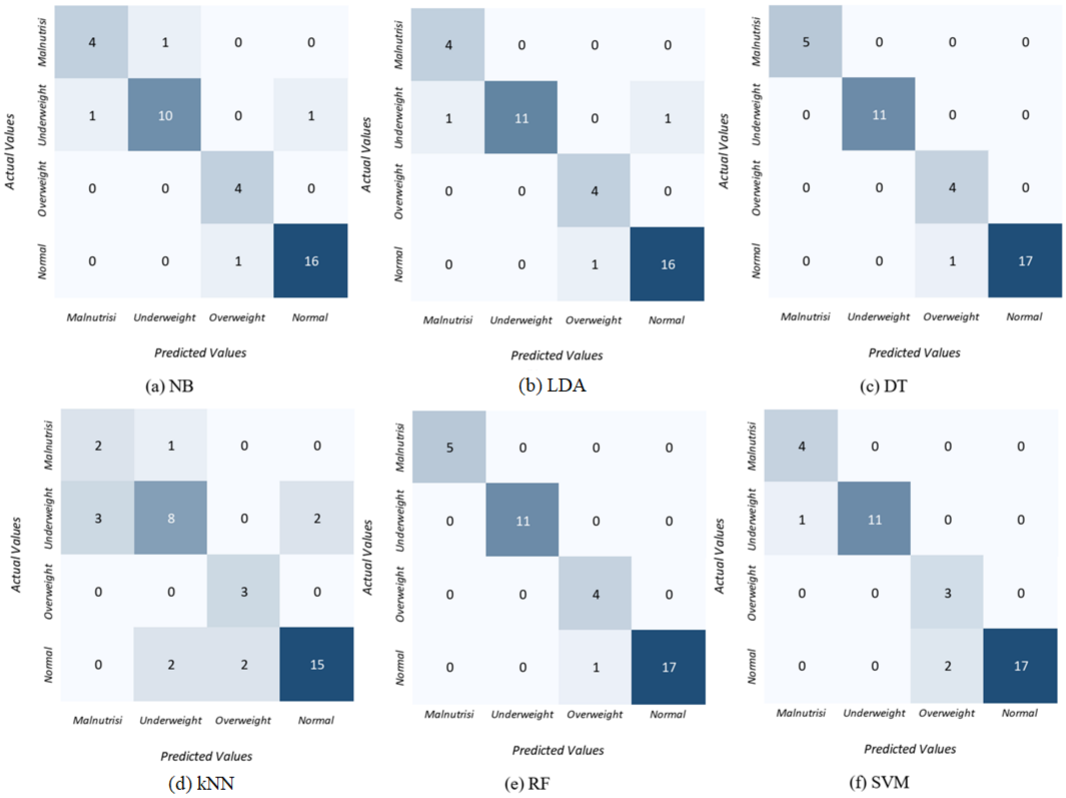


Figure 3: Confusion matrix of ML methods.

## 3.1 Dataset Analysis Results

The dataset contains 200 toddler data. There was no data reduction after it was verified from null, outliers, and duplication data. The data distribution of each variable was relatively even, as shown in Figure 2, except for the BMI variable, which was slightly skewed to the right. The Spearman rank correlation value between variables was low, except for the weight and BMI variables that significantly influenced the outcome variable (under-five status). The status plot also shows that the dataset is slightly unbalanced but quite good.

Furthermore, the training dataset was classified using ML methods (NB, LDA, DT, kNN, RF, SVM) for modeling. The best model of each ML method would be tested to measure its performance.

## 3.2 Confusion Matrix with k-Fold CV

Figure 3 shows the CM for each ML method with a 10-fold CV. The results of testing accuracy, sensitivity, and specificity metrics using (1), (2), and (3) on average are listed in Table 3. The results of CM testing with a 10-fold CV for all ML methods show that almost all ML methods have high accuracy, sensitivity, and specificity values of more than 0.81 except the kNN, which has an accuracy value of 0.7368 and sensitivity of 0.6524. Meanwhile, DT and RF are the best-performing ML methods.

Table 3: ML methods performance based on CM

| No. | Methods | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|
| 1. | NB | 0.8947 | 0.8626 | 0.9620 |
| 2. | LDA | 0.9211 | 0.8853 | 0.9696 |
| 3. | DT | 0.9737 | 0.9500 | 0.9881 |
| 4. | kNN | 0.7368 | 0.6524 | 0.8985 |
| 5. | RF | 0.9737 | 0.9500 | 0.9881 |
| 6. | SVM | 0.9609 | 0.8500 | 0.9670 |

## 3.3 AUC

The AUC measurement results in Figure 4 show that almost all ML methods have an outstanding level of discrimination except the kNN, which is at an excellent level. Meanwhile, the RF is the ML method with the highest AUC value.

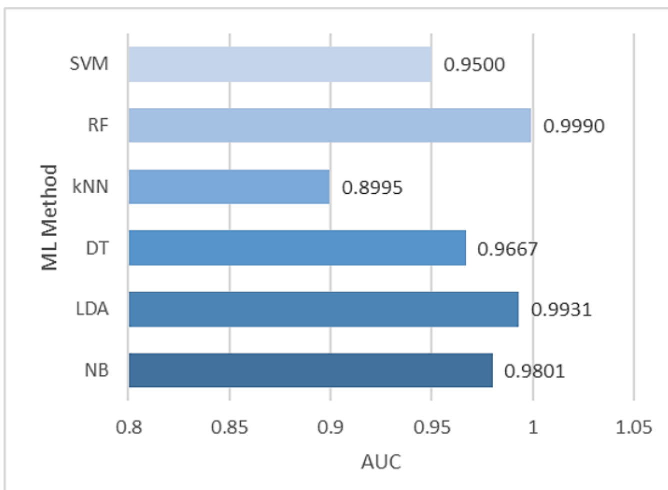


Figure 4: AUC values for the ML methods.

## 3.4 CKC

The reliability test results for all ML methods are based on the CKC value. Figure 5 shows that almost all ML methods have almost perfect reliability, with CKC values of more than 0.81, except the kNN has moderate reliability. Meanwhile, the DT and RF are the ML methods with the best reliability.

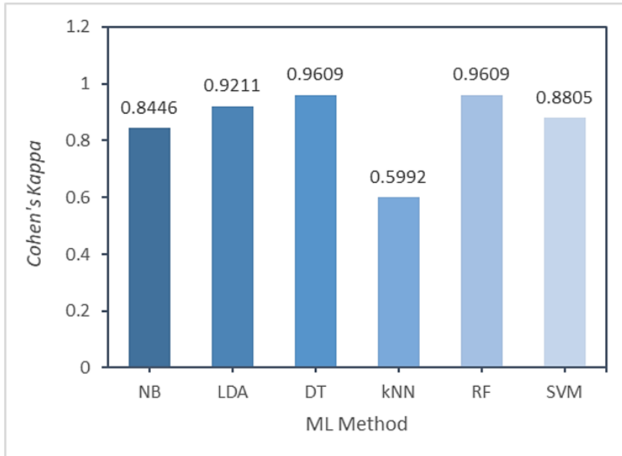


Figure 5: Reliability comparison of ML methods based on CKC.

# 4 Discussion

Test results based on accuracy, sensitivity, specificity, and reliability metrics show that RF and DT are ML methods with the best performance. The results of the RF and DT tests are similar to the limitations of the dataset scale and the number of variables. Hence, the RF decision combination is identical to the DT. However, test results based on AUC values show that RF performance outperforms DT and all other ML methods. Therefore, RF is the most recommended ML method for predicting toddlers' nutritional status. Whereas kNN is an ML method that is not recommended because it has the lowest performance based on all test results.

# 5 Conclusion

This research uses several ML methods, NB, LDA, DT, kNN, RF, and SVM, to predict toddlers' nutritional status based on age, BMI, weight, and length. Testing results using the CM with a 10-fold CV, AUC, and CKC show that almost all ML methods have high performance. However, the RF is the best-performing ML method.

In addition, the limitation of this research is that the dataset used is limited and only uses four variables (age, weight, height, and BMI) based on anthropometric standards in Indonesia. Future research can use datasets with a larger scale to increase the scope of the study and involve several other variables such as demographics and so on.

# Acknowledgments

Thank you to the Directorate General of Vocational Education, Ministry of Education, Culture, Research, and Technology for funding this research through the Applied Research-Matching Fund scheme in 2022 with contract number 316/PKS/D.D4/PPK.01.APTV/VII/2022.

## References

[1] Kementerian Kesehatan RI, Hasil Survei Status Gizi Indonesia (SSGI) 2022. Jakarta: Kementerian Kesehatan RI, 2023.

[2] UNICEF, "Child malnutrition," New York, 2023.

[3] E. R. Arumi, S. A. Subrata, and A. Rahmawati, "Implementation of naïve Bayes method for predictor prevalence level for malnutrition toddlers in Magelang City," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 2, pp. 201–207, 2023.

[4] M. Aryuni, E. Miranda, M. Kumbangsila, Richard, A. Y. Zakiyyah, A. V. D. Sano, and F. M. Bhatti, "Comparison of nutritional status prediction models of children under 5 years of age using supervised machine learning," in *3rd International Conference on Electronics, Biomedical Engineering, and Health Informatics*, 2023, pp. 265–277.

[5] R. Setiawan and A. Triayudi, "Klasifikasi status gizi balita menggunakan naïve Bayes dan k-nearest neighbor berbasis web," *J. Media Inform. Budidarma*, vol. 6, no. 2, pp. 777–785, 2022.

[6] A. Lasarudin, H. Gani, and M. Tomayahu, "Perbandingan metode naïve Bayes dan C4.5 klasifikasi status gizi bayi balita," *SPECTA J. Technol.*, vol. 6, no. 3, pp. 273–283, 2022.

[7] G. W. Kassie and D. L. Workie, "Determinants of under-nutrition among children under five years of age in Ethiopia," *BMC Public Health*, vol. 20, no. 1, pp. 1–11, Dec. 2020.

[8] S. Lonang and D. Normawati, "Klasifikasi Status Stunting Pada Balita Menggunakan K-Nearest Neighbor Dengan Feature Selection Backward Elimination," *J. Media Inform. Budidarma*, vol. 6, no. 1, pp. 49–56, 2022.

[9] A. Nazir, A. Akhyar, Y. Yusra, and E. Budianita, "Toddler nutritional status classification using C4.5 and particle swarm optimization," *Sci. J. Informatics*, vol. 9, no. 1, pp. 32–41, 2022.

[10] M. Ula, A. F. Ulva, M. Mauliza, M. A. Ali, and Y. R. Said, "Application of machine learning in determining the classification of children's nutrition with decision tree," *J. Tek. Inform.*, vol. 3, no. 5, pp. 1457–1465, 2022.

[11] K. Pinaryanto, R. A. Nugroho, and Y. Basilius, "Classification of toddler nutrition using C4.5 decision tree method," *Int. J. Appl. Sci. Smart Technol.*, vol. 3, no. 1, pp. 131–142, 2021.

[12] E. Ramon, A. Nazir, N. Novriyanto, Y. Yusra, and L. Oktavia, "Klasifikasi status gizi bayi posyandu Kecamatan Bangun Purba menggunakan algoritma support vector machine (SVM)," *J. Sist. Inf. dan Inform.*, vol. 5, no. 2, pp. 143–150, 2022.

[13] J. Pardede and H. Athifah, “Implementation of principal component analysis and learning vector quantization for classification of food nutrition status,” *JUITA J. Inform.*, vol. 10, no. 1, pp. 19–28, 2022.

[14] S. M. J. Rahman, N. A. M. F. Ahmed, M. M. Abedin, B. Ahammed, M. Ali, M .J. Rahman, and M. Maniruzzaman, “Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach,” *PLoS One*, vol. 16, no. 6, pp. 1–11, 2021.

[15] Z. Momand, P. Mongkolnam, P. Kositpanthavong, and J. H. Chan, “Data mining based prediction of malnutrition in Afghan children,” in *2020 12th International Conference on Knowledge and Smart Technology (KST)*, 2020, pp. 12–17.

[16] A. Talukder and B. Ahammed, “Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh,” *Nutrition*, vol. 78, pp. 1–22, 2020.

[17] F. H. Bitew, C. S. Sparks, and S. H. Nyarko, “Machine learning algorithms for predicting under nutrition among under-five children in Ethiopia,” *Public Health Nutr.*, vol. 25, no. 3, pp. 269–280, 2021.

[18] M. M. Shahriar, M. S. Iqubal, S. Mitra, and A. K. Das, “A deep learning approach to predict malnutrition status of 0-59 month’s older children in Bangladesh,” in *2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, 2019, pp. 145–149.

[19] S. A. Hemo and M. I. Rayhan, “Classification tree and random forest model to predict under-five malnutrition in Bangladesh,” *Biom Biostat Int J*, vol. 10, no. 3, pp. 116–123, 2021.

[20] Kaggle, “The baby nutrition dataset,” Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/mjalaluddinassuyuti/baby-nutrition-classification. [Accessed: 08-Jun-2023].

[21] Y. Chen, L. Li, W. Li, Q. Guo, Z. Du, and Z. Xu, “Fundamentals of neural networks,” in *AI Computing Systems*, Elsevier, 2024, pp. 17–51.

[22] S. Cahyani, R. Wiryasaputra, and R. Gustriansyah, “Identifikasi huruf kapital tulisan tangan menggunakan linear discriminant analysis dan Euclidean distance,” *J. Sist. Inf. Bisnis*, vol. 8, no. 1, pp. 57–67, Apr. 2018.

[23] A. Prakesakwa, A. Suryani, and R. Gustriansyah, “Sistem pendukung keputusan untuk subsidi biaya perbaikan kerusakan kontainer menggunakan naive Bayes,” *J. Teknol. dan Sist. Komput.*, vol. 7, no. 3, pp. 98–102, 2019.

[24] R. Gustriansyah, B. G. Ramadhan, N. Suhandi, A. Sanmorino, and J. R. Coyanda, “Determining employee eligibility in equalizing staffing status using the naïve Bayes,” *Int. J. Eng. Adv. Technol.*, vol. 9, no. 1, pp. 2309–2312, 2019.

[25] A. Sanmorino, “A study for DDOS attack classification method,” *J. Phys. Conf. Ser.*, vol. 1175, pp. 1–6, 2019.

[26] A. Sanmorino, R. Gustriansyah, and J. Alie, "DDoS attacks detection method using feature importance and support vector machine," *JUITA J. Inform.*, vol. 10, no. 2, pp. 167–171, 2022.

[27] S. Rafieyan, E. Vasheghani-Farahani, N. Baheiraei, and H. Keshavarz, "MLATE: Machine learning for predicting cell behavior on cardiac tissue engineering scaffolds," *Comput. Biol. Med.*, vol. 158, pp. 1–11, 2023.

[28] T. Kim and J.-S. Lee, "Maximizing AUC to learn weighted naive Bayes for imbalanced data classification," *Expert Syst. Appl.*, vol. 217, pp. 1–17, 2023.

[29] C. Yao, "Hearing loss classification via stationary wavelet entropy and cat swarm optimization," in *Cognitive Systems and Signal Processing in Image Processing*, Elsevier, pp. 203–221, 2022.