

CHAPTER 1 INTRODUCTION

1.1 Background

A child under the age of five, commonly referred to as a toddler, undergoes rapid brain development during this crucial period. Toddler nutrition plays a vital role, as this phase, known as the *golden age*, is critical for a child's growth and development. Proper nutrition during this period supports brain development, immune system function, and physical abilities, forming the foundation for a child's future health and learning capabilities. Adequate nutritional intake ensures that a child grows into an intelligent, sharp, and active individual.¹

Toddlers in their growth phase require adequate nutritional intake. During this period, cases of stunted growth, also known as *stunting*, remain prevalent in Indonesia. Stunting is a nutritional deficiency disorder commonly experienced by children in their early developmental years. According to the *Survei Status Gizi Indonesia* (SSGI), the stunting rate in Indonesia was recorded at 24.4% in 2021 and declined to 21.6% in 2022. Historically, the prevalence of stunting in Indonesia was significantly higher, reaching an average of 36.4% between 2005 and 2017. The incidence of stunting in Indonesia remains substantial, with other Southeast Asian countries also experiencing similar issues, such as Myanmar (35%), Vietnam (23%), and Thailand (16%).²

Anthropometry is the simplest and most cost-effective method for determining nutritional status. Anthropometric measurement refers to the assessment used to evaluate an individual's nutritional condition based on body dimensions. According to the Health Regulation of the Republic of Indonesia No. 2 of 2020 on Child Anthropometric Standards, there are six categories of

¹ Hamsir Saleh, Muh Faisal, and Rachmat Irawan Musa, "Klasifikasi Status Gizi Balita Menggunakan Metode K-Nearest Neighbor," *Simtek : Jurnal Sistem Informasi Dan Teknik Komputer* 4, no. 2 (2019): 120–26, <https://doi.org/10.51876/simtek.v4i2.60>.

² Achmad Aria Reza¹ and Muhammad Syaifur Rohman², "Prediction Stunting Analysis Using Random Forest Algorithm and Random Search Optimization," *Jite* 7, no. 2 (2024): 534–44, <http://ojs.uma.ac.id/index.php/jite>.

nutritional status.³ The commonly used anthropometric indices include weight-for-age (BB/U), weight-for-height (BB/TB), and height-for-age (TB/U). For children under 60 months, weight-for-age (BB/U) is typically used.⁴ The classification categories based on the Length-for-Age (PB/U) or Height-for-Age (TB/U) index are Normal, Short, Very Short, and Tall.

The method used in this study is data mining, specifically the Naïve Bayes classification algorithm. Naïve Bayes is a statistical classifier that can be used to predict the probability of class membership. It offers high accuracy and speed when applied to large databases. Moreover, Naïve Bayes is an algorithm that minimizes error rates compared to other classification methods.⁵

Jogorogo Community Health Centre (Puskesmas Jogorogo) collects nutritional status data of infants and young children from each village within the Jogorogo sub-district. Every February and August, the health centre experiences a surge in visitors as residents come for measurements, weighing, and other health checks. These months are thus referred to as "weighing months" due to the significant number of people attending for health assessments. Medical personnel at the health centre measure each child's weight and height, recording the data manually. They then calculate the Z-score (standard deviation) using the anthropometric standard table for assessing nutritional status. The collected data is processed and classified by village using an application provided by the Indonesian Ministry of Health (Kemenkes), known as *Sigizi Terpadu* or *E-PPBGM* (Electronic Community-Based Nutrition Recording and Reporting). However, the nutrition specialists at the health centre face challenges due to the absence of a feature in the application that classifies height-for-age (HFA). This functionality is essential to help parents determine whether their child's height is appropriate for their age.

³ ALVI FAJRIN LBS, "Klasifikasi Gizi Buruk Pada Balita Menggunakan Algoritma C5.0 (Studi Kasus : Dinas Kesehatan Aceh Timur)" (UNIVERSITAS MALIKUSSALEH LHOKSEUMAWE, 2024).

⁴ "Peraturan Menteri Kesehatan Republik Indonesia Nomor 2 Tahun 2020 Tentang Standar Antropometri Anak," 2020.

⁵ Widya Cholid Wahyudin, Fida Maisa Hana, and Agung Prihandono, "Prediksi Stunting Pada Balita Di Rumah Sakit Kota Semarang Menggunakan Naive Bayes" 2019 (2023): 32–36.

In Surah Al-A'raf (7:31).

يَا بَنِي آدَمَ خُذُوا زِينَتَكُمْ عِنْدَ كُلِّ مَسْجِدٍ وَكُلُوا وَشَرِبُوا وَلَا تُسْرِفُوا إِنَّهُ لَا يُحِبُّ

الْمُسْرِفِينَ (الأعراف: ٣١)

Translation: "O children of Adam! Wear your best attire at every place of worship, and eat and drink, but do not be excessive. Indeed, Allah does not like those who are excessive."

This verse teaches that Allah instructs humankind to always consume food and drink that is lawful, nutritious, beneficial, and enjoyable, while ensuring that it does not pose any harm to health and is not consumed excessively.⁶

A study conducted in 2023 by Nurainun, Elin Haerani, Fadhilah Syarifah, and Lola Oktavia focused on the classification of nutritional status based solely on the Weight-for-Height (BB/TB) index. The classification consisted of six categories: severely undernourished, undernourished, well-nourished, at risk of being overweight, overweight, and obese. The study found that the Naïve Bayes Classifier algorithm, combined with K-Fold Cross Validation, was able to classify nutritional status accurately. The highest recorded accuracy was 82.94% in the fifth iteration, while the lowest accuracy was 65.88% in the sixth iteration. The overall average accuracy was 75.47%, with a precision rate of 81.36% and a recall rate of 75.47%.

A study conducted in 2024 by Raya Srkandi, Dwi Hikmah Ramadhani, Muhammad Ikhwan, and Rizal Adi Saputra focused on classifying the nutritional status of toddlers based on age, weight, and height variables. In the classification process, the variables of age and weight were used with fuzzy sets including baduta (infants under two years old), batita (toddlers under three years old), pre-school, severely underweight, underweight, ideal, severely stunted, stunted, and normal. Following system testing with a total of 129 data points, 15 instances were misclassified, while 114 were correctly classified. Consequently, the achieved accuracy rate was 88.3%.

⁶ Veramia Ayu Windiyani, "Hubungan Asupan Energi, Aktivitas Fisik Dan Pengetahuan Gizi Seimbang Dengan Status Gizi Siswa Sma Negeri 8 Semarang," *Skripsi*, 2016, 1–23.

A study conducted in 2022 by Heru Budi Setiawan Gunawan Pria Utama focused on classifying the nutritional status of children under five. The research aimed to facilitate the determination of nutritional status using data mining techniques with the Naïve Bayes algorithm. The system was developed using PHP programming language and MySQL database. The study utilised a dataset of 551 records, employing Fold Cross-Validation with attributes including name, age, gender, weight, and nutritional status indices such as weight-for-age (BB/U), height-for-age (PB/U), and weight-for-height (BB/PB). The evaluation results using the Naïve Bayes method showed an accuracy rate of 72% for BB/U, 70.8% for PB/U, and 71.02% for BB/TB.

Therefore, the objective of this study is to facilitate the extraction of information regarding the classification of height based on age, enabling mothers with young children to determine whether their child's height is appropriate for their age.

1.2 Problem Formulation

The background of the problem outlined in the previous chapter presents several key issues, such as:

The processing of nutritional status data for toddlers at Puskesmas Jogorogo is currently conducted manually and classified using the *Sigizi Terpadu* or *E-PPBGM* application. However, the available application (*Sigizi Terpadu/E-PPBGM*) does not yet include a specific classification feature for Height-for-Age (TB/U) that can be directly utilised by medical personnel at the *puskesmas*. This means that healthcare workers still need to perform manual calculations before entering the data into the Ministry of Health's application.

1.3 Research Objective

Based on the issues formulated in the previous discussion, the objective of this study is:

To classify the nutritional status at *Puskesmas Jogorogo* using the Naïve Bayes Algorithm based on the Anthropometric Index, specifically the Height-for-Age (TB/U) category, obtained from Puskesmas Jogorogo.

1.4 Benefits

This research is expected to provide benefits to several parties, as follows:

1.4.1 For the Author

- a. Processing healthcare data and developing it into useful information for determining an accurate nutritional status.
- b. Identifying the best method for classifying the nutritional status threshold in toddlers using the Naïve Bayes algorithm..

1.4.2 For Readers

- a. Serving as a reference for research in Machine Learning and Data Mining.
- b. Providing knowledge about the digitalization of the healthcare sector, particularly in Data Mining.

1.4.3 For Society

- a. Contributing to community health monitoring at *Puskesmas Jogorogo*.
- b. Supporting the digitalization of the healthcare sector through the application of Machine Learning algorithms.

1.4.4 For the University

- a. Demonstrating that students can contribute to the healthcare sector with modern solutions.
- b. Providing evidence that students have mastered their field of study and successfully applied their knowledge to benefit society.

1.5 Scope of the Study

- a. The classification threshold for weight, height, and Body Mass Index (BMI) parameters is based on the Decree of the Minister of Health of the Republic of Indonesia No. 2 of 2020 concerning the nutritional assessment of toddlers using anthropometric standards.
- b. The classification of toddler nutritional status is divided into four categories based on the TB/U (Height-for-Age) or PB/U (Length-for-Age) index: severely stunted, stunted, normal, and tall.
- c. The research data is obtained from Puskesmas Jogorogo.

- d. The data used covers the period from 2023 to 2024 for children aged 0–5 years.
- e. The attributes used in the study are: Gender, Height, Age, and Z-score for TB/U or PB/U.
- f. The dataset consists of toddlers.

1.6 Systematic Discussion

This research will follow the writing structure as follows:

CHAPTER 1: INTRODUCTION

- 1.1 Background
- 1.2 Problem Formulation
- 1.3 Objectives
- 1.4 Benefits
- 1.5 Scope of the Study
- 1.6 Systematic Discussion

CHAPTER 2: LITERATURE REVIEW

- 2.1 Previous Research
 - 2.1.1 Nurainun, Elin Haerani, Fadhilah Syarifah, Lola Oktavia (2023)
 - 2.1.2 Raya Srkandi, Dwi Hikmah Ramadhani, Muhammad Ikhwan, Rizal Adi Saputra (2024)
 - 2.1.3 Heru Budi Setiawan Gunawan Pria Utama (2022)
 - 2.1.4 Annisa Maulana Majid, Rita Nuraeni, Abdul Halim Anshor (2020)
 - 2.1.5 Nuraisana, Sri Windarti Halawa, Muhammad Harun (2024)
- 2.2. Theoretical Foundation
 - 2.2.1 Classification
 - 2.2.2 Anthropometric Measurements
 - 2.2.3 Naïve Bayes
 - 2.2.4 Confusion Matrix

CHAPTER 3 RESEARCH METHODOLOGY

- 3.1 Research Time and Location

3.1.1 Time

3.1.2 Location

3.2 Research Tools and Materials

3.3 Research Phases

3.3.1 Business Understanding

3.3.2 Data Understanding

3.3.3 Data Preparation

3.3.4 Modelling

3.3.5 Evaluation

3.3.6 Deployment

CHAPTER 4 RESULTS AND DISCUSSION

4.1 Results

4.1.1 *Business Understanding*

4.1.2 *Data Understanding*

4.1.3 *Data Preparation*

4.1.4 Modelling

4.1.5 Evaluation

4.1.6 Deployment

4.2 Discussion

4.2.1 Nutritional Status Classification

4.2.2 Data Comparison between 2023 and 2024

CHAPTER 5 CONCLUSION

5.1 Conclusion

5.2 Recommendations

REFERENCES

APPENDICES

- a. Documentation of the research conducted at Puskesmas Jogorogo.
- b. Data Validation Sheet.