

**Implementing the K-Nearest Neighbors Algorithm on Toddler Nutritional Status Through
Comparison of Normalized Data
(Case Study: Puskesmas Minggir, Sleman)**

Yoga Tri Waskitho

*Program Studi Informatika, Medis Fakultas Sains dan Teknologi
Universitas Teknologi Yogyakarta
Jl. Ringroad Utara Jombor Sleman Yogyakarta
E-mail : yogatry036@gmail.com*

ABSTRACT

The application of machine learning models in predicting the nutritional status of toddlers has become a focal point of many studies, especially in efforts to reduce stunting rates. In Sleman Regency, the highest stunting rate is found in the Minggir District, with a prevalence of 8.6%. One crucial step in building a machine learning model is data preprocessing, which involves applying data normalization techniques to ensure the model's accuracy and reliability. However, the use of normalization techniques in predicting toddler nutritional status has not been widely explored. This study aims to build and compare the performance of classification models for predicting toddler nutritional status using the KNN algorithm, applying various normalization methods, including Min-Max, Z-score, and Decimal Scaling, as well as without normalization. The data used in this study comes from the Puskesmas Minggir, Sleman, consisting of 1,309 samples with 36 attributes. For the classification process, four input features were used: gender, age, weight, and height, along with three class labels: BB/U, TB/U, and BB/TB. The optimization of the $n_neighbors$ parameter was performed using Grid Search. The evaluation results showed that for the BB/U classification, the KNN model without normalization achieved the highest accuracy of 87%, followed by the KNN model with Z-score normalization, which had the second-best performance. In the TB/U classification, the KNN model without normalization also achieved the highest accuracy of 90%. In the BB/TB classification scheme, both the KNN model without normalization and the KNN model with Z-score normalization demonstrated equally strong performance, achieving an accuracy of 90% in each case.

Keywords: Nutritional Status Classification, K-Nearest Neighbors (KNN), Data Normalization, Grid Search