

# Analysis of Stunting Prediction for Toddlers in Bekasi Regency Using the K-Nearest Neighbors and Random Forest Algorithms

1<sup>st</sup> Kamelia Khoirunnisa

*School of Computing*

*Telkom University*

Bandung, Indonesia

kameliakhairunnisa@student.telkomuniversity.ac.id

2<sup>nd</sup> Putu Harry Gunawan

*School of Computing*

*Telkom University*

Bandung, Indonesia

phgunawan@telkomuniversity.ac.id

**Abstract**—Stunting, a condition where children are shorter than their age, is a serious nutritional issue in developing countries, including Indonesia. Research shows that Low Birth Weight (LBW) significantly affects children’s growth. In Bekasi Regency, the stunting prevalence remains high at 17%, with a target reduction to 14%. Achieving this target requires prevention efforts focused on improving nutrition and regularly monitoring child growth. In such monitoring, innovative evaluation techniques using Machine Learning (ML) are needed to predict stunting potential. This study aims to develop a predictive model for early detection of stunting risks in Bekasi Regency, using machine learning techniques to analyze Low Birth Weight (LBW) and Low Birth Length (LBL) factors, which could potentially be integrated into the local health monitoring system for preventive intervention. RF excels in handling complex features and identifying important predictors, while KNN is effective at recognizing local patterns. The results show that RF achieved the best performance with 99.22% accuracy and an F1-score of 96.94%, compared to KNN with 96.19% accuracy and an F1-score of 87.16%, highlighting RF’s greater stability and robustness over KNN in predicting stunting cases. This study is expected to provide an accurate predictive system that helps parents, health workers, and the government identify stunting potential early while also determining the appropriate ML algorithm for stunting case prediction in Indonesia. Future research is encouraged to test this model in other regions with different characteristics to ensure the generalizability and effectiveness of stunting prediction on a broader scale.

**Index Terms**—Stunting, Machine Learning, K-Nearest Neighbors, Random Forest

## I. INTRODUCTION

Stunting, a condition where children’s height is significantly lower than the average for their age, remains a critical nutritional challenge, particularly in developing and underdeveloped countries [1]. Research examining 26 stunting-related factors has identified Low Birth Weight (LBW) as a significant contributor that substantially impacts child growth and development [2]. Given LBW’s role as a causative factor, addressing stunting becomes increasingly crucial for preventing adverse effects on child development.

Indonesia’s commitment to reducing stunting rates to 14% by 2024, as stated by President Jokowi, has particular relevance for Bekasi Regency, which has demonstrated strong dedication to this cause. According to Acting Regent Dani Ramdan’s report on *bekasikab.go.id*, the region has already achieved a notable reduction in stunting rates from 21% to 17% between 2021 and 2022. However, reaching the 14% target requires sustained effort and enhanced collaboration among stakeholders.

This commitment aligns with WHO’s global target for stunting reduction by 2025. Indonesia, identified as one of 34 countries with high stunting prevalence, must contribute to WHO’s goal of reducing stunting rates by 40% by 2025 [3]. Bekasi Regency’s efforts directly support Indonesia’s commitment to achieving this target.

Machine Learning (ML) offers an innovative approach to stunting prevention through improved nutrition monitoring and child development assessment. ML techniques can predict potential stunting cases, enabling more proactive and targeted preventive measures [4].

Previous research in 2021 explored ML applications in addressing childhood stunting, specifically using the K-Nearest Neighbors classification algorithm. The study achieved optimal performance with an accuracy of 97.31% using specific data partition ratios [5]. A more recent study in February 2024 implementing Random Forest algorithms and Cross Validation for stunting prediction demonstrated good accuracy at 77.55% [6].

Another recent study in 2024 compared several ML algorithms for stunting prediction, evaluating Naive Bayes (NB), K-Nearest Neighbors (KNN), and Random Forest (RF). In this study, NB achieved an accuracy of 83.2%, KNN reached 84.8%, and RF achieved the highest accuracy at 87.75% [4]. A similar study also compared stunting prediction accuracy using KNN, RF, and Boosted KNN (BK) algorithms. In this study, KNN achieved an accuracy of 93.55%, RF was higher with an accuracy of 97.70%, and BK achieved the highest accuracy of 98.00% [7].

Although algorithms like K-Nearest Neighbors (KNN) and Random Forest (RF) have been used in stunting prediction, no study has yet compared the effectiveness of these algorithms in the context of stunting data using LBW and LBL factors, specifically in Bekasi Regency, which has different demographic and socio-economic characteristics compared to other regions. This study aims to conduct a comparative analysis between K-Nearest Neighbors and Random Forest algorithms using infant data based on LBW and LBL for stunting prediction. The results could assist healthcare providers and government officials in determining the most reliable predictive tool, ultimately supporting the implementation of more effective stunting prevention policies and interventions.