

PERBANDINGAN MODEL *NAÏVE BAYES*, *K-NEAREST NEIGHBOR* DAN *DECISION TREE C4.5* PADA KLASIFIKASI STATUS GIZI BALITA

ABSTRAK

Permasalahan terkait asupan gizi pada anak usia bawah lima tahun merupakan isu penting dalam pembangunan sumber daya manusia, terutama di wilayah terpencil yang memiliki keterbatasan akses terhadap informasi, pelayanan kesehatan, dan pemenuhan gizi. Penelitian ini bertujuan untuk membandingkan kinerja tiga algoritma klasifikasi yaitu *Naïve Bayes*, *K-Nearest Neighbor* (KNN), dan *Decision Tree C4.5* dalam mengklasifikasikan status gizi balita di Kecamatan Muara Dua dan Banyak Payed berdasarkan data dari aplikasi e-PPGBM tahun 2025. Status gizi diklasifikasikan menjadi dua kategori, yaitu gizi baik dan bermasalah gizi, dengan mempertimbangkan indikator antropometri BB/U, TB/U, dan BB/TB. Dataset yang dipakai terdiri dari 1.128 data balita dengan tujuh atribut, dan pengolahan data dilakukan menggunakan perangkat lunak *RapidMiner*. Evaluasi dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa algoritma KNN memiliki performa terbaik, dengan akurasi sebesar 76,92%, *precision* 86,40%, *recall* 63,91%, dan *F1-score* 73,47%. Algoritma *Naïve Bayes* unggul dalam aspek *recall* sebesar 98,82%, namun memiliki akurasi dan *precision* yang rendah. Sementara itu, *Decision Tree C4.5* menghasilkan performa yang seimbang tetapi masih di bawah KNN.

Kata Kunci: Status Gizi Balita, Naïve Bayes, K-Nearest Neighbor, Decision Tree C4.5, RapidMiner

A Comparative Study of Naïve Bayes, K-Nearest Neighbor, and Decision Tree C4.5 Algorithms for Classifying Nutritional Status in Toddlers

ABSTRACT

Nutritional problems among toddlers are a critical issue in the development of human resources, particularly in remote areas with limited access to information, healthcare services, and adequate nutrition. This study aims to compare the performance of three classification algorithms: Naïve Bayes, K-Nearest Neighbor (KNN), and Decision Tree C4.5 in classifying the nutritional status of toddlers in Muara Dua and Manyak Payed districts using data from the e-PPGBM application in 2025. Nutritional status is categorized into normal nutrition and nutritional problems, based on anthropometric indicators: weight-for-age (W/A), height-for-age (H/A), and weight-for-height (W/H). The dataset consists of 1,128 toddler records with seven attributes. Data processing was conducted using RapidMiner software, and model evaluation was carried out using accuracy, precision, recall, and F1-score metrics. The results show that the KNN algorithm outperforms the others with an accuracy of 76.92%, precision of 86.40%, recall of 63.91%, and F1-score of 73.47%. Although Naïve Bayes achieved the highest recall (98.82%), its accuracy and precision were lower. Meanwhile, Decision Tree C4.5 showed moderate performance but was still less effective than KNN.

Keywords: Nutritional Status, Toddlers, Naïve Bayes, K-Nearest Neighbor, Decision Tree C4.5, RapidMiner