

ABSTRACT

Micro, Small, and Medium Enterprises (MSMEs) in Indonesia face challenges in determining effective product strategies, especially when launching new products. The lack of utilization of historical sales data causes the decision-making process to be subjective and inefficient. This study aims to build a data-driven prediction model of product salability using the C4.5 decision tree algorithm, and to identify the most influential features contributing to sales potential. The research object is RumahBayiTAZ, an online-based store focusing on baby product sales. The dataset comprises seven main attributes, namely product name, quantity, product views, total visitors, number of new buyers, potential buyers, and total orders. Due to the class imbalance between popular and less popular products, the SMOTE technique is applied to produce a more balanced data distribution. The model is implemented using the ChefBoost library in Python and evaluated with 5-fold cross-validation. The results indicate that product views are the top priority feature, followed by new buyers and potential buyers as the most influential attributes for classification. The final model achieves an accuracy of 88.9% and provides interpretable decision rules that can guide business owners in designing product launch strategies. This research is expected to contribute practically to data-driven decision-making in the MSME sector.

Keywords: C4.5 algorithm, ChefBoost, product salability, feature priority, SMOTE, MSMEs.