

Sentiment Analysis of Laptop Product Reviews Using the XGBoost Algorithm on the Tokopedia Platform

Supervised by Raditya Arief Pratama, S.Kom., M.Eng.

Syafrizal Wd Mahendra
*Informatics Engineering Study Program
Department of Information Technology*

ABSTRACT

Indonesia's laptop market has grown rapidly in line with global trends recording sales of up to 348.8 million units in 2021, supported by BPS data indicating that 18.24 million Indonesians already own a computer or laptop. This growth has driven a surge in online transactions on e-commerce platforms such as Tokopedia, generating a massive and unstructured volume of product reviews. An information overload phenomenon emerges as a direct consequence, hindering consumers from objectively evaluating products, particularly given that 52% of online shopping preferences are driven by the availability of customer testimonials. This system helps users choose the right laptop by automatically analyzing thousands of customer reviews, so users don't have to read them one by one to find out a product's pros and cons. The methodology integrates web Scraping for data collection, text preprocessing, TF-IDF feature weighting, and three optimization techniques like (Synthetic Minority Over-sampling Technique (SMOTE), Chi-Square, and Grid Search). The model with best hyperparameter value achieves an accuracy of 73%, a Macro F1-Score of 0.67. Functional validation through blackbox testing using the equivalence partitioning technique recorded a 100% success rate across 8 Test Cases. The integration of SMOTE and hyperparameter optimization proves effective in minimizing majority class bias, enabling the system to deliver objective and structured sentiment insights for prospective laptop buyers on the Tokopedia platform.

Key words: XGBoost, Sentiment Analysis, Tokopedia, SMOTE, Grid Search, Laptop.