

ABSTRAK

Pertumbuhan platform e-commerce menyebabkan semakin banyaknya produk yang ditawarkan sehingga diperlukan metode yang dapat membantu mengidentifikasi tingkat permintaan produk secara efektif. Penelitian ini bertujuan untuk membandingkan kinerja algoritma Random Forest dan Support Vector Machine (SVM) dalam mengklasifikasikan permintaan produk pada platform e-commerce Tokopedia ke dalam tiga kategori, yaitu Low, Mid, dan High. Data diproses melalui tahapan prapemrosesan, ekstraksi fitur, dan hyperparameter tuning sebelum dilakukan evaluasi menggunakan metrik akurasi, macro F1-score, confusion matrix, cross-validation, dan learning curve. Hasil penelitian menunjukkan bahwa SVM dengan kernel Radial Basis Function (RBF) merupakan model dengan performa terbaik, dengan akurasi sebesar 85,47%, macro F1-score sebesar 82,57%, dan cross-validation macro F1-score sebesar 81,65% (standard deviation: 0,0059). Model SVM menunjukkan performa terbaik pada kelas Low dan terendah pada kelas High yang dipengaruhi oleh ketidakseimbangan distribusi data dan kemiripan karakteristik produk antar kelas. Meskipun SVM memiliki kecenderungan overfitting yang lebih tinggi, model tetap menunjukkan kemampuan generalisasi yang baik. Oleh karena itu, SVM dengan kernel RBF direkomendasikan sebagai model yang paling efektif untuk klasifikasi permintaan produk pada penelitian ini.

Kata kunci: klasifikasi permintaan produk, Random Forest, Support Vector Machine, e-commerce, Tokopedia, machine learning.

ABSTRACT

The growth of e-commerce platforms has resulted in an increasing number of products being offered, creating a need for methods that can effectively identify product demand levels. This study aims to compare the performance of the Random Forest and Support Vector Machine (SVM) algorithms in classifying product demand on the Tokopedia e-commerce platform into three categories: Low, Mid, and High. The data were processed through preprocessing, feature extraction, and hyperparameter tuning stages before being evaluated using accuracy, macro F1-score, confusion matrix, cross-validation, and learning curve metrics. The results showed that SVM with the Radial Basis Function (RBF) kernel achieved the best performance, with an accuracy of 85.47%, a macro F1-score of 82.57%, and a cross-validation macro F1-score of 81.65% with a standard deviation of 0.0059. The SVM model achieved the best performance on the Low class and the lowest performance on the High class, which was influenced by the imbalanced data distribution and similarities in product characteristics across classes. Although SVM showed a higher tendency toward overfitting, the model still demonstrated good generalization ability. Therefore, SVM with the RBF kernel is recommended as the most effective model for product demand classification in this study.

Keywords: product demand classification, Random Forest, Support Vector Machine, e-commerce, Tokopedia, machine learning.

