

ABSTRAK

Pertumbuhan pesat *e-commerce* di Indonesia telah mendorong peningkatan volume ulasan pengguna yang sulit dikelola secara manual. Shopee sebagai salah satu platform *e-commerce* terbesar menerima ribuan ulasan harian dengan keragaman bahasa dan ekspresi yang kompleks. Tantangan dalam mengelola data ulasan ini menjadi dasar perlunya penerapan teknologi *analisis sentimen* yang efisien dan akurat. Penelitian ini berfokus pada optimalisasi *analisis sentimen* terhadap ulasan pengguna Shopee dengan membandingkan tiga metode populer: *Naive Bayes*, *Support Vector Machine* (SVM), dan *Bidirectional Encoder Representations from Transformers* (BERT).

Penelitian ini menggunakan pendekatan *eksperimental* dengan *dataset* ulasan Shopee berbahasa Indonesia yang diperoleh dari platform publik. Data melalui proses *preprocessing* seperti *tokenisasi*, *normalisasi*, *stopword removal*, dan *stemming*. Selanjutnya, dilakukan pelatihan model menggunakan tiga metode klasifikasi: *Naive Bayes* dengan representasi *TF-IDF*, *SVM* dengan optimasi parameter melalui *GridSearchCV*, dan *BERT* dengan *fine-tuning* pada model *IndoBERT*. Evaluasi performa model dilakukan menggunakan metrik *akurasi*, *presisi*, *recall*, dan *F1-score*, serta validasi dengan teknik *train-test split*.

Hasil penelitian menunjukkan bahwa model *BERT* memiliki performa tertinggi dengan *akurasi* mencapai 90%, diikuti oleh *SVM* sebesar 85%, dan *Naive Bayes* sebesar 78%. *BERT* unggul dalam memahami konteks dan nuansa bahasa alami, terutama dalam menangani struktur kalimat kompleks dalam ulasan. Meskipun *SVM* juga menunjukkan kinerja yang baik dalam klasifikasi, metode ini memerlukan proses optimasi parameter yang lebih intensif. *Naive Bayes*, meskipun sederhana dan cepat, memiliki keterbatasan dalam menangani konteks linguistik yang kompleks, namun tetap relevan sebagai *baseline*.

Penelitian ini menyimpulkan bahwa model *BERT* merupakan metode paling efektif untuk *analisis sentimen* ulasan Shopee dalam konteks bahasa Indonesia. Implikasi dari temuan ini membuka peluang penerapan *BERT* secara lebih luas dalam sistem rekomendasi dan pengelolaan *feedback* pengguna di platform *e-commerce*. Penelitian selanjutnya disarankan untuk mengeksplorasi penggabungan model atau pendekatan *ensemble* guna meningkatkan *akurasi* dan efisiensi klasifikasi sentimen secara berkelanjutan.

Kata kunci: *analisis sentimen, Shopee, Naive Bayes, Support Vector Machine, BERT, e-commerce*



ABSTRACT

The rapid growth of e-commerce in Indonesia has driven a surge in user reviews, which are difficult to manage manually. Shopee, as one of the largest e-commerce platforms, receives thousands of reviews daily with diverse language and expressions. The challenge in managing this data underlines the need for an efficient and accurate sentiment analysis approach. This study focuses on optimizing sentiment analysis of Shopee user reviews by comparing three popular methods: Naive Bayes, Support Vector Machine (SVM), and Bidirectional Encoder Representations from Transformers (BERT).

This research adopts an experimental approach using an Indonesian-language Shopee review dataset obtained from a public platform. The data underwent preprocessing steps such as tokenization, normalization, stopwords removal, and stemming. The models were trained using three classification techniques: Naive Bayes with TF-IDF representation, SVM with parameter tuning via GridSearchCV, and BERT with fine-tuning on the IndoBERT model. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics, and validated with a train-test split strategy.

The results showed that the BERT model achieved the highest performance with an accuracy of 90%, followed by SVM at 85%, and Naive Bayes at 78%. BERT excelled in understanding the contextual nuances of natural language, especially in handling complex sentence structures. Although SVM also performed well, it required more intensive parameter tuning. Naive Bayes, while simple and fast, struggled with linguistic complexity but remained useful as a baseline method.

This study concludes that BERT is the most effective method for sentiment analysis of Shopee reviews in the Indonesian language context. These findings suggest broader applications of BERT in recommendation systems and user feedback management on e-commerce platforms. Future research is encouraged to explore model integration or ensemble approaches to further improve classification accuracy and efficiency.

Keywords: sentiment analysis, Shopee, Naive Bayes, Support Vector Machine, BERT, e-commerce