

ABSTRAK

Tujuan Penelitian ini untuk melakukan perbandingan kinerja antara dua algoritma yang sering digunakan, yakni *Naive Bayes* dan *K Nearest Neighbors (KNN)*, dalam analisis tingkat kepuasan pelanggan berdasarkan ulasan produk di platform *e-commerce* Tokopedia. Ulasan produk yang diberikan oleh pelanggan berperan penting dalam menentukan reputasi dan kualitas produk, serta memengaruhi keputusan pembelian calon pelanggan yang lain. Metodologi penelitian melibatkan pengumpulan data

ulasan produk dari Tokopedia, yang kemudian melalui tahapan pra-pemrosesan seperti pembersihan teks, tokenisasi, dan stemming. selanjutnya, ulasan tersebut diklasifikasikan kedalam dua kategori sentimen, yakni positif dan negatif dengan memanfaatkan algoritma *Naive Bayes* dan *KNN*. Kinerja kedua algoritma dievaluasi berdasarkan metrik *accuracy*, *recall*, *f1-score* dan *precision*. Penelitian ini menunjukkan bahwa algoritma *Naive Bayes* mempunyai *accuracy* lebih tinggi dalam mengkategorikan ulasan produk dibandingkan dengan *KNN*. Analisis ini memberikan pengetahuan yang berguna bagi perusahaan untuk memilih algoritma yang tepat dalam analisis sentimen ulasan produk, untuk meningkatkan kepuasan pelanggan. Penelitian ini bertujuan untuk memberikan kontribusi dalam memperluas pemahaman mengenai keunggulan dan keterbatasan masing-masing algoritma dalam analisis sentimen, serta memberikan rekomendasi bagi peneliti dalam mengimplementasikan algoritma yang tepat untuk mengukur tingkat kepuasan pelanggan.

Keywords: *Sentiment Analysis, Naïve Bayes, K Nearest Neighbors (KNN), Tokopedia*

ABSTRACT

This study conducts a comparative performance analysis of two widely utilized classification algorithms, *Naive Bayes* and *K-Nearest Neighbors (KNN)*, in the context of customer satisfaction analysis based on product reviews from the Tokopedia e-commerce platform. Customer-generated reviews serve as a critical factor in shaping product reputation and perceived quality, while also influencing the purchasing behavior of prospective buyers. The methodology encompasses data collection of product reviews from Tokopedia, followed by a comprehensive preprocessing pipeline, including text cleaning, tokenization, and stemming. The processed reviews are then categorized into two sentiment classes-positive and negative-employing both *Naive Bayes* and *KNN* algorithms. The performance of these algorithms is evaluated using standard classification metrics: *accuracy*, *recall*, *F1-score* dan *precision*. Empirical results demonstrate that *Naive Bayes* yields superior accuracy in classifying product sentiments compared to *KNN*. This research offers practical insights for e-commerce businesses in selecting suitable machine learning techniques for sentiment analysis to better understand customer feedback and enhance satisfaction. Moreover, the study contributes to the academic discourse by highlighting the strengths and limitations of each algorithm, and provides recommendations for future research in developing effective sentiment classification frameworks for customer satisfaction measurement..

Keywords: *Sentiment Analysis, Naive Bayes, K Nearest Neighbors (KNN), Tokopedia*