

INTISARI

Penelitian ini bertujuan untuk membandingkan kinerja algoritma *Support Vector Machine (SVM)*, *Naïve Bayes (NB)*, dan *K-Nearest Neighbors (KNN)* dalam analisis sentimen ulasan aplikasi *Netflix*. Data yang digunakan berupa 10.000 ulasan pengguna dari *Google Play Store* yang telah melalui proses *preprocessing teks*, termasuk *case folding*, *tokenisasi*, *stopword removal*, dan *stemming*. Representasi fitur menggunakan metode **TF-IDF**. Untuk mengatasi ketidakseimbangan data, penelitian ini menerapkan tiga teknik resampling, yaitu *SMOTE*, *SMOTETomek*, dan *SMOTEENN*. Model dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *cross-validation*. Hasil penelitian menunjukkan bahwa algoritma *SVM* dengan teknik *SMOTEENN* memberikan kinerja terbaik dengan *accuracy* 98,04%, diikuti oleh *KNN* dengan 97,65%, dan *Naïve Bayes* dengan 91,76%. Teknik *SMOTEENN* terbukti lebih efektif dibandingkan metode lain dalam meningkatkan *accuracy* model karena tidak hanya menambahkan data sintesis tetapi juga menghilangkan data mayoritas yang ambigu. Hasil *Cross Validation* menunjukkan bahwa *SVM* memiliki performa paling stabil dalam mengklasifikasikan sentimen ulasan pengguna. Berdasarkan temuan ini, disimpulkan bahwa *SVM* dengan teknik *SMOTEENN* adalah kombinasi terbaik untuk analisis sentimen ulasan aplikasi *Netflix*, karena mampu menangani ketidakseimbangan data secara optimal dan memberikan hasil klasifikasi yang lebih akurat dibandingkan algoritma lain.

Kata kunci: *Analisis Sentimen, Support Vector Machine, Naïve Bayes, K-Nearest Neighbors, Netflix*.

ABSTRACT

This study aims to compare the performance of the Support Vector Machine (SVM), Naïve Bayes (NB), and K-Nearest Neighbors (KNN) algorithms in analyzing Netflix app review sentiment. The data used are 10,000 user reviews from the Google Play Store that have gone through text preprocessing, including case folding, tokenization, stopwords removal, and stemming. Feature representation uses the TF-IDF method. To overcome data imbalance, this study applies three resampling techniques, namely SMOTE, SMOTETomek, and SMOTEENN. The model is evaluated using accuracy, precision, recall, F1-score, and Cross Validation metrics. The results show that the SVM algorithm with the SMOTEENN technique provides the best performance with an accuracy of 98.04%, followed by KNN with 97.65%, and Naïve Bayes with 91.76%. The SMOTEENN technique is proven to be more effective than other methods in improving model accuracy because it not only adds synthetic data but also eliminates the majority of ambiguous data. The Cross Validation results show that SVM has the most stable performance in classifying user review sentiment. Based on these findings, it is concluded that SVM with SMOTEENN technique is the best combination for Netflix app review sentiment analysis, because it is able to handle data imbalance optimally and provide more accurate classification results than other algorithms.

Keywords: Sentiment Analysis, Support Vector Machine, Naïve Bayes, K-Nearest Neighbors, Netflix.