

INTISARI

Tesis ini berjudul “Perbandingan Kinerja Algoritma *Machine Learning* Klasik dalam Analisis Sentimen Ulasan Pengguna Spotify Berbahasa Indonesia di Google Play Store”. Penelitian ini dilakukan karena ulasan pengguna di Google Play Store berjumlah besar, bersifat tidak terstruktur, dan mengandung bahasa informal sehingga sulit dianalisis secara manual. Tujuan penelitian adalah membandingkan kinerja beberapa algoritma *machine learning* klasik dalam mengklasifikasikan sentimen ulasan menjadi tiga kelas (positif, netral, dan negatif). Data penelitian diperoleh melalui proses *scraping* sebanyak 20.000 ulasan Spotify berbahasa Indonesia, dengan pelabelan sentimen berbasis rating. Metode analisis mencakup *preprocessing* teks, representasi fitur menggunakan TF-IDF, serta pengujian lima algoritma klasifikasi, yaitu Naive Bayes, Logistic Regression, Support Vector Machine (SVM), Random Forest, dan XGBoost, termasuk perbandingan fitur TF-IDF berbasis kata (*word n-gram*) dan berbasis karakter (*char_wb*).

Hasil pengujian menunjukkan bahwa pemilihan algoritma dan representasi fitur memengaruhi performa klasifikasi, terutama karena distribusi kelas tidak seimbang (*class imbalance*). Oleh karena itu, evaluasi dilakukan tidak hanya menggunakan *accuracy*, tetapi juga *macro-F1*. Konfigurasi terbaik pada penelitian ini adalah XGBoost dengan TF-IDF Char (*char_wb* 3–5) yang menghasilkan *test_accuracy* 0,8818 dan *test_macro-F1* 0,5703. Analisis *confusion matrix* dan *classification report* menunjukkan performa sangat baik pada kelas positif dan cukup baik pada kelas negatif, namun masih lemah pada kelas netral.

Kesimpulan penelitian ini menyatakan bahwa pendekatan *machine learning* klasik berbasis TF-IDF efektif untuk analisis sentimen ulasan Spotify berbahasa Indonesia, dengan XGBoost + TF-IDF Char sebagai konfigurasi yang paling seimbang pada data uji. Namun, peningkatan kemampuan model dalam mengenali kelas netral secara konsisten masih menjadi tantangan utama pada data yang tidak seimbang.

Kata kunci: *sentiment analysis*, Spotify, *TF-IDF*, *Support Vector Machine*, ulasan Google Play Store

ABSTRACT

This thesis, entitled “Comparative Evaluation of Classical Machine Learning Algorithms for Sentiment Analysis of Indonesian-Language Spotify User Reviews on the Google Play Store,” addresses the challenge of extracting insights from large-scale, unstructured, and informal app reviews. The study aims to compare the performance of classical machine learning classifiers in categorizing Spotify reviews into three sentiment classes (positive, neutral, and negative). A total of 20,000 Indonesian-language reviews were collected through web scraping from the Google Play Store, and sentiment labels were derived from user ratings. The analysis pipeline includes text preprocessing, TF-IDF feature representation, and the evaluation of five classification algorithms: Naive Bayes, Logistic Regression, Support Vector Machine (SVM), Random Forest, and XGBoost. In addition, two feature settings are compared, namely word-level TF-IDF (word n-grams) and character-level TF-IDF (char_wb).

The experimental results indicate that both algorithm selection and feature representation affect classification performance, particularly under class imbalance conditions. Therefore, evaluation is not limited to accuracy but also considers macro-F1 to reflect balanced performance across classes. The best-performing configuration is XGBoost with character-level TF-IDF (char_wb 3–5), achieving a test accuracy of 0.8818 and a macro-F1 score of 0.5703. Further analysis using the confusion matrix and classification report shows strong performance in the positive class and adequate performance in the negative class, while the neutral class remains the most challenging to detect.

In conclusion, classical TF-IDF-based machine learning approaches are effective for Indonesian-language Spotify review sentiment analysis, with XGBoost using character-level TF-IDF providing the most balanced performance on the test set. However, improving the consistency of neutral-class detection remains a key challenge, especially in the presence of class imbalance.

Keywords: sentiment analysis, Spotify, TF-IDF, Support Vector Machine, Google Play Store reviews