ANALISIS SENTIMEN ULASAN APLIKASI LETTERBOXD DI GOOGLE PLAY MENGGUNAKAN METODE NAÏVE BAYES DAN INFORMATION GAIN

Fernaldi Anggadha

Abstrak

Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna aplikasi Letterboxd pada platform Google Play Store dengan menggunakan metode klasifikasi Naïve Bayes serta seleksi fitur Information Gain. Dalam era digital, ulasan pengguna menjadi indikator penting untuk menilai kualitas dan kinerja suatu aplikasi. Naïve Bayes digunakan karena efisien dalam mengklasifikasikan teks, sedangkan Information Gain dimanfaatkan untuk menyaring fitur-fitur yang paling relevan guna meningkatkan performa model. Penelitian ini menerapkan pendekatan kuantitatif dengan data ulasan yang dikumpulkan melalui API google-play-scraper dari Januari hingga Mei 2025. Proses praproses data mencakup case folding, pembersihan teks, tokenisasi, penghilangan stopword, dan lemmatization. Data kemudian dibagi dengan rasio 80:20 untuk pelatihan dan pengujian. Hasil evaluasi menunjukkan bahwa model Naïve Bayes tanpa seleksi IG (*Information Gain*) mampu mencapai akurasi sebesar 82,5% dengan nilai precision dan recall masing-masing sebesar 83% dan 84% serta model Naïve Bayes dengan seleksi IG mampu mencapai akurasi 83,5% dengan nilai prescision dan recall masing-masing sebesar 84% dan 85%. Selain itu, dilakukan analisis lanjutan dengan memetakan hasil sentimen negatif ke dalam tujuh karakteristik kualitas perangkat lunak berdasarkan standar ISO/IEC 25010 menggunakan pendekatan rule-based. Temuan ini diharapkan dapat menjadi dasar bagi pengembang Letterboxd dalam meningkatkan aspek kualitas aplikasi berdasarkan persepsi dan pengalaman pengguna.

Kata kunci: analisis sentimen, Naïve Bayes, Information Gain, ISO 25010, Google Play, Letterboxd.

SENTIMENT ANALYSIS OF LETTERBOXD APPLICATION REVIEWS ON GOOGLE PLAY USING NAÏVE BAYES AND INFORMATION GAIN

Fernaldi Anggadha

Abstract

This research aims to analyze user sentiment from reviews of the Letterboxd application on the Google Play Store using the Naïve Bayes classification method and Information Gain feature selection. In the digital era, user reviews serve as a critical indicator for evaluating application quality and performance. Naïve Bayes was chosen for its efficiency in text classification, while Information Gain was applied to filter the most relevant features to enhance model performance. A quantitative approach was employed, with review data collected via the google-play-scraper API between January and May 2025. Text preprocessing steps included case folding, cleaning, tokenization, stopword removal, and lemmatization. The data was then split using an 80:20 ratio for training and testing. Evaluation results showed that the Naïve Bayes model achieved 82.5% accuracy with precision and recall scores of 83% and 84%, respectively, meanwhile, combined with Information Gain, Naïve Bayes achieved an accuracy of 83.5%, with a precision of 84% and a recall of 85%.. Furthermore, a follow-up analysis was conducted by mapping negative sentiment results into seven software quality characteristics based on the ISO/IEC 25010 standard using a rule-based classification approach. These findings are expected to assist Letterboxd developers in enhancing application quality based on user sentiment and experience.

Keywords: *sentiment analysis, Naïve Bayes, Information Gain, ISO 25010, Google Play, Letterboxd.*