

INTISARI

Penelitian ini bertujuan untuk mengevaluasi penerapan algoritma Support Vector Machine (SVM) dalam melakukan analisis sentimen terhadap review aplikasi LinkedIn. Tujuan utama penelitian adalah untuk memahami cara penerapan algoritma SVM dalam menganalisis sentimen pada teks review aplikasi LinkedIn. Penelitian ini menggunakan dataset sebanyak 2.535 review dari pengguna di Indonesia, dengan klasifikasi sentimen positif dan negatif sebagai fokus analisis. Metode penelitian ini melibatkan proses klasifikasi sentimen dengan memanfaatkan nilai rata-rata (means) dari skor review. Dalam hal ini, nilai skor kurang dari 3 dianggap sebagai teks negatif (0), sedangkan nilai skor lebih besar dari 3 dianggap sebagai teks positif (1). Hasil penyederhanaan kelas menunjukkan bahwa terdapat 1.035 teks negatif dan 1.500 teks positif. Dalam tahap preprocessing data meliputi tahap tokenize dan stemming, pembobotan data menggunakan fungsi Tf-Idf lalu mengimplementasikan dengan algoritma Support Vector Machine menggunakan fungsi kernel Linear. Penerapan SVM dengan data latih 80% dan data uji 20% menggunakan fungsi kernel Linear serta parameter gamma dan Cost (C). Evaluasi model dilakukan menggunakan confusion matrix, dan hasilnya menunjukkan bahwa model memiliki kinerja yang baik. F1-Score mencapai 89,91%, recall sebesar 88,36%, precision sebesar 91,53%, dan akurasi sebesar 87,57%.

Kata Kunci: Analisis Sentimen, Support Vector Machine, LinkedIn, kernel Linear, Klasifikasi, Confusion Matrix.

ABSTRACT

This research aims to evaluate the implementation of the Support Vector Machine (SVM) algorithm in sentiment analysis of LinkedIn application reviews. The main objective is to comprehend the application of the SVM algorithm in analyzing sentiment in LinkedIn application reviews. The study employs a dataset comprising 2.535 reviews from users in Indonesia, with positive and negative sentiment classifications as the focus of analysis. The research methodology involves sentiment classification using the mean values of review scores. In this context, scores below 3 are considered negative text (0), while scores exceeding 3 are deemed positive text (1). The simplification of classes reveals 1.035 negative texts and 1.500 positive texts. The data preprocessing phase includes tokenization and stemming, with data weighting utilizing the Tf-Idf function implemented through the Support Vector Machine algorithm with a kernel Linear. SVM implementation involves an 80% training data and 20% testing data split, utilizing the kernel Linear. Model evaluation is conducted using a confusion matrix, demonstrating a robust performance of the model. The F1-Score reaches 89.91%, recall stands at 88.36%, precision is 91.53%, and accuracy is 87.57%.

Keywords: Sentiment Analysis, Support Vector Machine, LinkedIn, kernel Linear, Classification, Confusion Matrix.