

ABSTRAK

Perkembangan aplikasi pembelajaran digital menghasilkan semakin banyak ulasan pengguna yang dapat dimanfaatkan untuk mengevaluasi kualitas layanan melalui analisis sentimen. Namun, performa klasifikasi sentimen tidak hanya dipengaruhi oleh algoritma yang digunakan, tetapi juga oleh metode pelabelan data. Penelitian ini bertujuan membandingkan pengaruh pendekatan *lexicon-based* menggunakan Indonesian Sentiment Lexicon (InSet) dengan pelabelan manual menggunakan metode *majority voting* terhadap performa klasifikasi sentimen ulasan pengguna Ruangguru menggunakan algoritma *Support Vector Machine* (SVM). Data penelitian berupa 5.000 ulasan pengguna yang diperoleh dari Google Play Store. Tahapan penelitian meliputi *text preprocessing*, pelabelan vembalivt menggunakan kedua pendekatan, ekstraksi fitur menggunakan *Term Frequency–Inverse Document Frequency* (TF-IDF), penyeimbangan data menggunakan *Synthetic Minority Oversampling Technique* (SMOTE), serta klasifikasi menggunakan SVM. Evaluasi model dilakukan menggunakan *confusion matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *Macro F1-score*. Hasil penelitian menunjukkan bahwa model SVM dengan pelabelan *lexicon-based* memberikan performa yang lebih baik dibandingkan pelabelan manual, dengan nilai *accuracy* sebesar 84,88%, *Macro Precision* 0,75, *Macro Recall* 0,78, dan *Macro F1-score* 0,76, sedangkan pelabelan manual menghasilkan *accuracy* sebesar 77,17%, *Macro Precision* 0,67, *Macro Recall* 0,70, dan *Macro F1-score* 0,68. Temuan ini menunjukkan bahwa metode pelabelan memengaruhi performa klasifikasi sentimen, sehingga pendekatan *lexicon-based* menggunakan InSet lebih efektif dalam menghasilkan model klasifikasi sentimen ulasan pengguna Ruangguru.

Kata kunci: Analisis Sentimen, *Support vector Machine*, *Lexicon-Based*, SMOTE

ABSTRACT

The rapid growth of digital learning applications has generated an increasing number of user reviews that can be utilized to evaluate service quality through sentiment analysis. However, the performance of sentiment classification is influenced not only by the classification algorithm but also by the data viembaivit approach. This study aims to compare the effect of a lexicon-based approach using the Indonesian Sentiment Lexicon (InSet) with manual viembaivit using the majority voting method on the performance of sentiment classification of Ruangguru user reviews using the Support Vector Machine (SVM) algorithm. The dataset consisted of 5,000 user reviews collected from the Google Play Store. The research process included text preprocessing, sentiment viembaivit using both approaches, feature extraction using Term Frequency–Inverse Document Frequency (TF-IDF), data balancing using the Synthetic Minority Oversampling Technique (SMOTE), and sentiment classification using SVM. Model performance was evaluated using a confusion matrix with Accuracy, Precision, Recall, and Macro F1-score as the evaluation metrics. The results showed that the SVM model trained using lexicon-based viembaivit achieved better performance than the model trained using manual viembaivit, obtaining an accuracy of 84.88%, a Macro Precision of 0.75, a Macro Recall of 0.78, and a Macro F1-score of 0.76. In comparison, the manually viembaiviti model achieved an accuracy of 77.17%, a Macro Precision of 0.67, a Macro Recall of 0.70, and a Macro F1-score of 0.68. These findings indicate that the viembaivit approach affects the performance of sentiment classification, demonstrating that the InSet-based lexicon approach is more effective for developing sentiment classification models for Ruangguru user reviews.

Keywords :Sentiment Analysis, Support Vector Machine, Lexicon-Based, SMOTE