

ABSTRAK

Ulasan pengguna pada aplikasi MyBCA di Google Play Store mengandung berbagai opini yang dapat digunakan untuk mengetahui persepsi pengguna terhadap aplikasi. Namun, jumlah ulasan yang terus bertambah menyebabkan proses analisis secara manual menjadi kurang efisien sehingga diperlukan metode analisis sentimen yang mampu mengolah data secara otomatis. Penelitian ini bertujuan untuk mengimplementasikan model IndoBERT sebagai metode pelabelan sentimen otomatis serta menerapkan algoritma Support Vector Machine (SVM) untuk klasifikasi sentimen ulasan aplikasi MyBCA di Google Play Store. Data yang digunakan berupa 5.000 ulasan pengguna berbahasa Indonesia yang diperoleh melalui teknik web scraping. Tahapan penelitian meliputi preprocessing data berupa case folding, cleaning text, dan normalization text, kemudian pelabelan sentimen menggunakan model pra-latih IndoBERT yang mengklasifikasikan ulasan ke dalam kategori positif, netral, dan negatif. Dataset selanjutnya dibagi menjadi data latih sebesar 80% dan data uji sebesar 20%. Proses ekstraksi fitur dilakukan menggunakan TF-IDF dan n-gram untuk mengubah teks menjadi representasi numerik sebelum dilakukan klasifikasi menggunakan algoritma SVM. Hasil penelitian menunjukkan bahwa distribusi sentimen didominasi oleh sentimen negatif sebesar 57,7%, diikuti sentimen positif sebesar 33,3%, dan sentimen netral sebesar 9%. Berdasarkan hasil evaluasi model, diperoleh accuracy sebesar 90,4%, precision sebesar 89,63%, recall sebesar 90,4%, dan F1-score sebesar 89,91%. Hasil tersebut menunjukkan bahwa kombinasi IndoBERT, TF-IDF, dan SVM mampu mengklasifikasikan sentimen ulasan aplikasi MyBCA dengan performa yang baik, khususnya pada kelas sentimen positif dan negatif.

Kata kunci: analisis sentimen, IndoBERT, SVM, TF-IDF, MyBCA.

ABSTRACT

User reviews on the MyBCA application available on Google Play Store contain various opinions that can be utilized to understand users' perceptions of the application. However, the continuously increasing number of reviews makes manual analysis inefficient, thereby requiring an automated sentiment analysis approach. This study aims to implement IndoBERT as an automatic sentiment labeling method and Support Vector Machine (SVM) as a sentiment classification algorithm for MyBCA user reviews on Google Play Store. The dataset consisted of 5,000 Indonesian-language reviews collected through web scraping. The research stages included data preprocessing, namely case folding, cleaning text, and normalization text, followed by sentiment labeling using a pre-trained IndoBERT model that classified reviews into positive, neutral, and negative sentiments. The dataset was then divided into 80% training data and 20% testing data. Feature extraction was performed using TF-IDF and n-gram techniques to transform text into numerical representations before classification using SVM. The results showed that negative sentiment dominated the reviews with 57.7%, followed by positive sentiment with 33.3%, and neutral sentiment with 9.0%. The evaluation results demonstrated that the SVM model achieved an accuracy of 90.4%, precision of 89.63%, recall of 90.4%, and F1-score of 89.91%. These findings indicate that the combination of IndoBERT, TF-IDF, and SVM is capable of effectively classifying sentiment in MyBCA user reviews, particularly for positive and negative sentiment classes.

Keywords: sentiment analysis, IndoBERT, SVM, TF-IDF, MyBCA.