

INTISARI

ANALISIS SENTIMEN BERBASIS ASPEK PADA ULASAN PENGGUNA APLIKASI KAI ACCESS MENGGUNAKAN *SUPPORT VECTOR MACHINE* DAN *RANDOM OVERSAMPLING*

Oleh

Huda Mustakim

17/412641/PA/17960

Tingginya penggunaan internet telah mempengaruhi berbagai hal. Salah satunya adalah dengan diluncurkannya aplikasi KAI Access oleh PT Kereta Api Indonesia (PT KAI) sebagai sistem pemesanan tiket online. Sayangnya, banyak ditemukan ulasan negatif di *Google Play Store* yang harus diperhatikan pihak pengembang. Sudah ada penelitian pada ulasan tersebut menggunakan metode klasifikasi *Support Vector Machine* (SVM), namun hasil klasifikasi hanya mampu mengklasifikasi sentimen pada ulasan secara umum, sehingga sentimen dan informasi terkait aspek penting seperti aspek-aspek *usability* aplikasi belum diketahui. Selain itu, belum adanya penanganan *imbalance data* menyebabkan akurasi prediksi pada kelas minoritas menjadi rendah.

Pada penelitian ini dilakukan analisis sentimen berbasis aspek untuk mengetahui sentimen pengguna aplikasi KAI Access terhadap aspek *usability* aplikasi yaitu *learnability*, *efficiency*, *errors*, dan *satisfaction*, menggunakan metode SVM, dengan penerapan *Random Oversampling* untuk mengatasi *imbalance data*, serta beberapa tahap *preprocessing* data dengan ekstraksi fitur menggunakan TF-IDF. Untuk mendapatkan hasil klasifikasi yang lebih baik, juga dilakukan *hyperparameter tuning* pada model SVM menggunakan *GridSearch CV*. Sementara itu, data penelitian ini diambil dari Google Play Store.

Dari hasil penelitian, berhasil dibuat model klasifikasi dari metode SVM yang mampu mendeteksi sentimen ulasan di setiap aspek aplikasi, dengan performa yang cukup baik, yakni dengan rata-rata skor dari seluruh model klasifikasi dari hasil pengujian yang memiliki skor akurasi 91,24%, dan *f1-score* 71,42%. Penerapan teknik *oversampling* juga berhasil meningkatkan *f1-score* dan kemampuan model klasifikasi pada kelas minoritas. Sementara itu, hasil analisis menunjukkan bahwa mayoritas sentimen bernilai negatif di seluruh aspek aplikasi terutama pada aspek *errors*, yang menandakan tingginya tingkat kesalahan sistem.

Kata kunci: analisis sentimen berbasis aspek, *support vector machine*, *imbalance data*, *oversampling*, kai access

ABSTRACT

ASPECT BASED SENTIMENT ANALYSIS OF KAI ACCESS APPLICATION USER REVIEWS USING SUPPORT VECTOR MACHINE AND RANDOM OVERSAMPLING

by

Huda Mustakim

17/412641/PA/17960

The high use of the internet has affected various things. One of them is PT Kereta Api Indonesia (PT KAI) which has launched KAI Access application as online ticket booking system. However, since its launch, many negative reviews have been found in Google Play Store that developers must pay attention to. There has been research on the review using classification method Support Vector Machine (SVM), but the system created is only able to classify sentiment on review data in general, so that sentiment and information regarding importance aspects related to the application are not yet known. In addition, the absence of handling imbalance data causes prediction accuracy in the minority class is poor.

In this study, an aspect-based sentiment analysis was carried out to determine the sentiments of users of the KAI Access application on the usability aspects of the application, namely learnability, efficiency, errors, and satisfaction, using the SVM method, with the application of Random Oversampling to overcome data imbalance, as well as several stages of preprocessing data with feature extraction. using TF-IDF. To get better classification results, hyperparameter tuning is also performed on the SVM model using GridSearch CV. Meanwhile, this research data was taken from the Google Play Store.

From the results of the study, a classification model from the SVM method was successfully created and able to detect sentiment on each review for every application aspect, with a fairly good performance, with average score of all classification models from the test results has accuracy score of 91.24%, and f1-score 71.42%. The use of oversampling technique also successfully enhance the f1-score and the model's ability in classifying the minority class. Meanwhile, the results of the analysis show that the majority of sentiments are negative in all application aspects especially on errors aspect, which indicates a high level of system error.

Keywords: aspect based sentiment analysis, support vector machine, imbalance data, oversampling, kai access