

ABSTRACT

Online transportation is a widely used digital service that generates a large number of user reviews on the Google Play Store. These reviews contain not only general sentiment but also opinions on specific aspects, such as service, drivers, applications, and price/cost. Therefore, Aspect-Based Sentiment Analysis (ABSA) is needed to provide a more specific understanding of user perceptions. This study aims to apply ABSA to user reviews of Grab, Gojek, and Maxim using Logistic Regression with TF-IDF feature extraction. The research data were obtained from a Kaggle dataset containing 300,000 reviews, which were filtered to 58,867 Indonesian-language reviews. Manual annotation samples were selected using Stratified Random Sampling with Equal Allocation, resulting in 2,001 reviews that were further split into text units. The final labels were used as ground truth, with Cohen's Kappa values of 0.7187 for aspect annotation and 0.8887 for polarity annotation. In the modeling stage, three scenarios were compared: *Baseline* Logistic Regression, *Class Weight Balanced*, and SMOTE. The *testing* results show that Logistic Regression with *Class Weight Balanced* achieved the best performance, with an *Accuracy* of 80.28%, a *Macro F1-score* of 72.29%, and a *Weighted F1-score* of 80.07%. The aspect-based sentiment analysis results indicate that Maxim performed well in the driver and price/cost aspects, Grab showed a relatively balanced sentiment distribution, while Gojek obtained high negative sentiment in several main aspects. This study shows that ABSA can provide more detailed information about user perceptions of *online* transportation services.

Keywords: Aspect-Based Sentiment Analysis, Logistic Regression, TF-IDF, *Online* Transportation, User Reviews

ABSTRAK

Transportasi *online* merupakan layanan digital yang banyak digunakan masyarakat dan menghasilkan banyak ulasan pengguna pada Google Play Store. Ulasan tersebut tidak hanya memuat sentimen umum, tetapi juga opini terhadap aspek tertentu, seperti layanan, pengemudi, aplikasi, dan harga/biaya. Oleh karena itu, diperlukan pendekatan *Aspect-Based Sentiment Analysis* (ABSA) untuk memperoleh gambaran persepsi pengguna secara lebih spesifik. Penelitian ini bertujuan untuk menerapkan ABSA pada ulasan pengguna aplikasi Grab, Gojek, dan Maxim menggunakan *Logistic Regression* dengan ekstraksi fitur TF-IDF. Data penelitian diperoleh dari dataset Kaggle yang berisi 300.000 ulasan, kemudian melalui proses penyaringan hingga diperoleh 58.867 ulasan berbahasa Indonesia. Sampel anotasi manual ditentukan menggunakan *Stratified Random Sampling* dengan *Equal Allocation* sebanyak 2.001 ulasan yang kemudian dipecah menjadi unit teks. Label akhir digunakan sebagai *ground truth* dengan nilai Cohen's Kappa sebesar 0,7187 untuk aspek dan 0,8887 untuk polaritas. Pada tahap pemodelan, dilakukan perbandingan tiga skenario, yaitu *Baseline Logistic Regression*, *Class Weight Balanced*, dan SMOTE. Hasil pengujian menunjukkan bahwa *Logistic Regression* dengan *Class Weight Balanced* memperoleh performa terbaik dengan *Accuracy* sebesar 80,28%, *Macro F1-score* sebesar 72,29%, dan *Weighted F1-score* sebesar 80,07%. Hasil analisis menunjukkan bahwa Maxim unggul pada aspek pengemudi dan harga/biaya, Grab memiliki distribusi sentimen yang relatif seimbang, sedangkan Gojek memperoleh sentimen negatif tinggi pada beberapa aspek utama. Penelitian ini menunjukkan bahwa ABSA dapat memberikan informasi yang lebih rinci mengenai persepsi pengguna terhadap layanan transportasi *online*.

Kata kunci: Analisis Sentimen Berbasis Aspek, Logistic Regression, TF-IDF, Transportasi *Online*, Ulasan Pengguna