

ABSTRAK

TIARA SONYA. Penerapan Metode *IndoBERT* untuk Analisis Sentimen Ulasan Pengguna Aplikasi Ferizy di *Google Play Store*. Dibimbing oleh MARTALELI BETTIZA dan NOLAN EFRANDA.

Aplikasi Ferizy sebagai layanan pemesanan tiket transportasi laut milik PT ASDP Indonesia Ferry mengalami peningkatan jumlah pengguna yang pesat, sehingga menghasilkan volume ulasan yang besar pada Google Play Store dan menjadikan evaluasi sentimen secara manual tidak lagi efisien untuk data berskala besar. Penelitian ini menerapkan model *IndoBERT* untuk mengklasifikasikan sentimen positif dan negatif terhadap 8.000 ulasan hasil web scraping. Setelah melalui tahap praproses berupa cleaning, case folding, dan normalisasi serta penghapusan label netral, diperoleh 6.574 data yang didominasi oleh sentimen negatif. Dataset kemudian dibagi dengan rasio 70:10:20 untuk data latih, validasi, dan uji, lalu diseimbangkan menggunakan Random Undersampling (RUS) sehingga menghasilkan 4.964 data dengan distribusi kelas yang seragam. Hyperparameter tuning dilakukan melalui grid search terhadap enam kombinasi learning rate dan batch size, dengan model terbaik dipilih berdasarkan nilai F1-Macro paling tinggi dari data validasi. Hasil pengujian pada data testing menunjukkan model *IndoBERT* pada (Skenario 1 Data Tidak Seimbang) mencapai akurasi sebesar 94,45%, precision 94,25%, recall 93,92%, dan F1-Macro sebesar 94,08%, sedangkan model *IndoBERT* (Skenario 2 Data Seimbang) memperoleh akurasi, precision, recall, dan F1-Macro sebesar 93,36%. Penelitian ini menyimpulkan bahwa model *IndoBERT* efektif dan stabil diterapkan untuk menganalisis sentimen ulasan dalam bahasa Indonesia dengan data berskala besar.

Kata kunci: Analisis Sentimen, *IndoBERT*, Ferizy, Grid Search, Random Undersampling.

ABSTRACT

TIARA SONYA. *Implementation of the IndoBERT Method for Sentiment Analysis of Ferizy Application User Reviews on the Google Play Store. Supervised by MARTALELI BETTIZA and NOLAN EFRANDA.*

The Ferizy application, a marine transportation ticketing service owned by PT ASDP Indonesia Ferry, experienced rapid user growth, generating a vast volume of Google Play Store reviews that rendered manual sentiment evaluation inefficient. This study implemented the IndoBERT model to classify positive and negative sentiments across 8,000 web-scraped reviews. Following preprocessing including cleaning, case folding, normalization, and the removal of neutral labels, 6,574 instances remained, dominated by negative sentiment. The dataset was split into training, validation, and test sets at a 70:10:20 ratio, and subsequently balanced using Random Undersampling (RUS), resulting in 4,964 instances with a uniform class distribution. Hyperparameter tuning via grid search evaluated six combinations of learning rates and batch sizes, selecting the optimal model based on the highest validation F1-Macro score. Test results demonstrated that IndoBERT on Scenario 1 (Imbalanced Data) achieved an accuracy of 94.45%, precision of 94.25%, recall of 93.92%, and an F1-Macro score of 94.08%. Meanwhile, IndoBERT on Scenario 2 (Balanced Data) achieved 93.36% across accuracy, precision, recall, and F1-Macro. This study concludes that the IndoBERT model is highly effective and stable when implemented for sentiment analysis of Indonesian-language reviews on a large scale.

Keywords: Sentiment Analysis, IndoBERT, Ferizy, Grid Search, Random Undersampling.