ABSTRAK

Perkembangan layanan financial technology (fintech), khususnya pinjaman online, mendorong meningkatnya jumlah ulasan pengguna pada Google Play Store sebagai media untuk menyampaikan pengalaman, kepuasan, maupun keluhan terhadap layanan yang digunakan. Ulasan tersebut mengandung informasi yang bermanfaat bagi penyedia layanan untuk mengevaluasi kualitas layanan, namun karakteristik bahasa yang tidak baku, penggunaan singkatan, istilah gaul, serta campuran bahasa Indonesia dan bahasa asing menjadi tantangan dalam proses analisis sentimen. Penelitian ini bertujuan mengimplementasikan serta menganalisis kinerja model Multilingual Bidirectional Encoder Representations from Transformers (mBERT) dalam mengklasifikasikan sentimen ulasan pengguna aplikasi pinjaman online di Indonesia. Metode yang digunakan adalah SEMMA (Sample, Explore, Modify, Model, Assess). Dataset diperoleh melalui proses scraping ulasan pengguna pada tiga aplikasi, yaitu Kredivo, Julo, dan Shopee PayLater. Tahap preprocessing meliputi pembersihan teks, normalisasi, pelabelan sentimen berdasarkan rating, tokenisasi menggunakan tokenizer mBERT, serta pembagian dataset dengan dua skenario, yaitu 80:10:10 dan 70:15:15. Model dilatih menggunakan arsitektur bert-base-multilingual-cased dengan learning rate 2×10^{-5}, batch size 8, dan 3 epoch. Evaluasi dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model mBERT mampu mengklasifikasikan sentimen dengan baik pada ketiga dataset. Kinerja terbaik diperoleh pada dataset Kredivo dengan skenario 70:15:15 yang menghasilkan accuracy 92,17%, precision 90,61%, recall 92,17%, dan F1-score 91,38%. Hasil tersebut menunjukkan bahwa mBERT memiliki kinerja yang baik dalam klasifikasi sentimen ulasan pengguna aplikasi pinjaman online berbahasa Indonesia dan dapat dimanfaatkan sebagai pendekatan untuk membantu memahami opini pengguna terhadap layanan fintech.

Kata Kunci: Analisis Sentimen, Google Play Store, *Multilingual BERT* (mBERT), *Natural Language Processing*, Pinjaman *Online*, SEMMA.

ABSTRACT

The growth of financial technology (fintech) services, particularly online lending, has led to an increase in user reviews on the Google Play Store, serving as a platform for users to share their experiences, satisfaction levels, and complaints. While these reviews contain valuable information for service providers to evaluate service quality, the analysis of sentiment is challenged by characteristics such as non-standard language, abbreviations, slang, and the mixing of Indonesian with foreign languages. This study aims to implement and analyze the performance of the Multilingual Bidirectional Encoder Representations from Transformers (mBERT) model in classifying the sentiment of user reviews for online lending applications in Indonesia. The SEMMA (Sample, Explore, Modify, Model, Assess) methodology was employed. The dataset was obtained by scraping user reviews from three applications: Kredivo, Julo, and Shopee PayLater. Preprocessing steps included text cleaning, normalization, sentiment labeling based on ratings, tokenization using the mBERT tokenizer, and dataset splitting under two scenarios: 80:10:10 and 70:15:15. The model was trained using the `bert-base-multilingual-cased` architecture with a learning rate of 2×10^{-5}, a batch size of 8, and 3 epochs. Performance was evaluated using accuracy, precision, recall, and F1-score metrics. The results demonstrate that the mBERT model effectively classifies sentiment across all three datasets. The best performance was achieved with the Kredivo dataset using the 70:15:15 split, yielding an accuracy of 92.17%, precision of 90.61%, recall of 92.17%, and an F1-score of 91.38%. These findings indicate that mBERT performs well in classifying the sentiment of Indonesian-language user reviews for online lending apps and can serve as a valuable approach for understanding user opinions regarding fintech services.

Keywords: Google Play Store, Multilingual BERT (mBERT), Natural Language, Online Lending, Processing, SEMMA, Sentiment Analysis.