

ABSTRACT

Gavrila Louise Tumanggor (01082210011)

ANALISIS SENTIMEN BERBASIS ASPEK: KLASIFIKASI ULASAN PRODUK KECANTIKAN DENGAN INDOBERT

(xiv + 130 pages: 29 figures; 21 tables, 1 appendices)

In the digital era, consumer reviews of products on e-commerce platforms have become a vital source of information for producers, sellers, and other consumers. However, extracting meaningful insights from the vast and diverse array of reviews requires an in-depth analytical approach. A key challenge in analyzing reviews is understanding user sentiment toward specific aspects, such as product quality, seller service, and delivery performance, rather than merely evaluating overall opinions. To address this, an Aspect-Based Sentiment Analysis (ABSA) method is essential, enabling the extraction of relevant aspects from reviews and the accurate classification of sentiments associated with each aspect.

This study proposes the implementation of the IndoBERT model to integrate the tasks of Aspect Extraction (AE) and Aspect Sentiment Classification (ASC) for beauty product reviews in Indonesian. IndoBERT was selected for its exceptional ability to comprehend the context of the Indonesian language. The research dataset consists of 10,000 product reviews categorized into negative, neutral, and positive sentiments based on review scores. The model is fine-tuned using hyperparameter configurations such as batch size and the number of epochs to enhance accuracy. The research stages include data preprocessing, aspect and sentiment labeling, model training, and performance evaluation. The expected outcome is a model capable of accurately extracting review aspects and classifying sentiments with a high level of confidence.

The model implementation achieved an accuracy of 86% and an F1-score of 0.857, with a balanced accuracy of 88%. It demonstrated precise predictions for aspects such as "product" (confidence score 0.98), "seller" (1.00), and "delivery" (0.86). Testing revealed the model could process data at a rate of 33.173 samples per second and complete training in 1.5072 seconds per iteration. These results confirm the model's effectiveness in identifying aspects and sentiments in reviews, significantly contributing to aspect-based sentiment analysis in the Indonesian language while supporting improvements in product service quality and marketing strategies.

Reference: 48 (2018 – 2024)

ABSTRAK

Gavrila Louise Tumanggor (01082210011)

ANALISIS SENTIMEN BERBASIS ASPEK: KLASIFIKASI ULASAN PRODUK KECANTIKAN DENGAN INDOBERT

(xiv + 130 pages: 29 figures; 21 tables, 1 appendices)

Dalam era digital, ulasan konsumen terhadap produk di platform *e-commerce* menjadi sumber informasi penting bagi produsen, penjual, dan konsumen lainnya. Namun, penggalan informasi dari ulasan yang sangat banyak dan beragam membutuhkan pendekatan analisis yang mendalam. Tantangan utama dalam menganalisis ulasan adalah memahami sentimen pengguna terhadap aspek spesifik, seperti kualitas produk, pelayanan penjual, dan pengiriman, bukan hanya sekadar opini secara keseluruhan. Untuk itu, dibutuhkan metode analisis sentimen berbasis aspek (*Aspect-Based Sentiment Analysis/ABSA*) yang mampu mengekstraksi aspek-aspek relevan dalam ulasan dan mengklasifikasikan sentimen yang terkandung pada setiap aspek secara akurat.

Penelitian ini mengusulkan penerapan model IndoBERT untuk menggabungkan tugas *Aspect Extraction* (AE) dan *Aspect Sentiment Classification* (ASC) pada ulasan produk kecantikan berbahasa Indonesia. IndoBERT dipilih karena keunggulannya dalam memahami konteks bahasa Indonesia secara mendalam. *Dataset* penelitian mencakup 10.000 ulasan produk dengan kategori sentimen negatif, netral, dan positif berdasarkan skor ulasan. Model dioptimalkan menggunakan teknik *fine-tuning* dengan konfigurasi parameter seperti ukuran batch dan jumlah epoch yang disesuaikan untuk meningkatkan akurasi. Tahapan penelitian meliputi praproses data, pelabelan aspek dan sentimen, pelatihan model, dan evaluasi performa. Hasil akhir yang diharapkan adalah model yang mampu mengekstraksi aspek-aspek ulasan secara akurat dan memberikan klasifikasi sentimen dengan tingkat kepercayaan yang tinggi.

Implementasi model menghasilkan akurasi 86% dan *F1-score* 0.857, dengan *balanced accuracy* mencapai 88%. Model juga menunjukkan kemampuan prediksi yang akurat pada aspek seperti "produk" (*confidence score* 0.98), "penjual" (1.00), dan "pengiriman" (0.86). Pengujian menunjukkan model dapat memproses data dengan kecepatan 33.173 sampel per detik dan menyelesaikan pelatihan dalam waktu 1.5072 detik per iterasi. Dengan hasil ini, model terbukti efektif dalam mengidentifikasi aspek dan sentimen ulasan, memberikan kontribusi penting bagi analisis sentimen berbasis aspek dalam bahasa Indonesia serta mendukung peningkatan kualitas layanan produk dan strategi pemasaran.

Referensi: 48 (2018 – 2024)