

INTISARI

PENINGKATAN PERFORMA *ASPECT-BASED SENTIMENT ANALYSIS* PADA *DATASET AWARE* MENGGUNAKAN RTT DAN H-SUM BERT

Oleh

Nadia Hasna Azzahra

21/477869/PA/20712

Performa rendah pada *dataset AWARE* dalam tugas *Aspect-Based Sentiment Analysis* (ABSA) menunjukkan dua tantangan utama, yaitu ketidakseimbangan data dan kompleksitas ABSA yang membutuhkan representasi semantik yang kaya. Untuk menjawab tantangan tersebut, penelitian ini mengukur kombinasi teknik augmentasi *Round-Trip Translation* (RTT) untuk mengatasi ketidakseimbangan data dan arsitektur H-SUM BERT sebagai *classifier* untuk meningkatkan representasi semantik. Evaluasi dilakukan pada tugas *Aspect Category Classification* (ACC) dan *Aspect Sentiment Classification* (ASC) menggunakan metrik mikro-F1 dan akurasi. Penelitian juga meninjau dampak *text cleaning* serta membandingkan efisiensi H-SUM BERT terhadap BERT *base*.

Hasil menunjukkan bahwa kombinasi RTT dan H-SUM BERT tanpa *text cleaning* menghasilkan performa terbaik. Pada tugas ACC, skor mikro-F1 mencapai 72,72 untuk domain *Productivity*, 75,00 untuk *Social Networking*, dan 43,94 untuk *Games*. Sementara pada tugas ASC, akurasi yang diperoleh masing-masing sebesar 88,87%, 86,38%, dan 75,00%. Meskipun terjadi peningkatan performa, penggunaan H-SUM BERT meningkatkan waktu pelatihan sebesar 28% dan waktu inferensi sebesar 34% dibandingkan BERT *base*.

Kata Kunci: H-SUM BERT, *Round-Trip Translation*, *Aspect-Based Sentiment Analysis*, Klasifikasi Teks, *AWARE Dataset*, Data Ulasan Aplikasi.

ABSTRACT

PERFORMANCE IMPROVEMENT OF ASPECT-BASED SENTIMENT ANALYSIS ON THE AWARE DATASET USING RTT AND H-SUM BERT

By

Nadia Hasna Azzahra

21/477869/PA/20712

Low performance on the AWARE dataset in the task of Aspect-Based Sentiment Analysis indicates two main challenges, data imbalance and the complexity of ABSA which requires deep semantic representation. To address these issues, this study evaluates the combination of Round Trip Translation as a data augmentation technique to handle data imbalance and H-SUM BERT architecture as a classifier to enhance semantic representation. The evaluation was conducted on the Aspect Category Classification (ACC) and Aspect Sentiment Classification (ASC) tasks using micro-F1 and accuracy metrics. This study also examines the impact of text cleaning and compares the efficiency of H-SUM BERT with BERT base.

Experimental results show that the combination of RTT and H-SUM BERT without text cleaning achieved the best performance. In the ACC task, the model achieved a micro-F1 score of 72.72 for the Productivity domain, 75.00 for Social Networking, and 43.94 for Games. Meanwhile, in the ASC task, the accuracies obtained were 88.87%, 86.38%, and 75.00%, respectively. Despite the improved performance, the use of H-SUM BERT incurs a 28% increase in training time and a 34% increase in inference time compared to BERT base.

Keywords: H-SUM BERT, *Round-Trip Translation*, *Aspect-Based Sentiment Analysis*, *Text Classification*, *AWARE Dataset*, *Application Review Data*.