

INTISARI

Analisis sentimen ulasan buku berbahasa Indonesia memerlukan model yang tidak hanya akurat tetapi juga efisien untuk mendukung sistem real-time. Penelitian ini mengevaluasi hasil optimasi IndoBERT menggunakan Multi-Objective Bayesian Optimization (MOBO) guna meningkatkan macro-F1 score dan menurunkan p95 inference latency. Dataset terdiri dari 3.269 ulasan berlabel positif dan negatif dari website Zahira Media Publisher yang dibagi secara stratified (70% latih, 15% validasi, 15% uji). Model IndoBERT (indobenchmark/indobert-base-p1) dilatih sebagai baseline, lalu dioptimasi menggunakan Optuna melalui Tree-structured Parzen Estimator (TPE) dan Random Sampler dalam skema multi-objektif. Evaluasi menggunakan macro-F1 score dan p95 inference latency.

Model baseline memperoleh macro-F1 sebesar 0,8715 dan p95 latency 25,5060 ms, sedangkan model hasil optimasi mencapai macro-F1 sebesar 0,9977 dan p95 latency 11,8410 ms. Hasil ini menunjukkan MOBO meningkatkan akurasi sekaligus efisiensi inferensi secara signifikan, sehingga konfigurasi IndoBERT hasil optimasi layak diterapkan pada sistem analisis sentimen real-time di Zahira Media Publisher.

Kata kunci: *IndoBERT, Bayesian Optimization, Multi-Objective Optimization, Sentiment Analysis, Optuna*

ABSTRACT

Sentiment analysis of Indonesian-language book reviews requires a model that is not only accurate but also efficient to support real-time systems. This study evaluates the optimization results of IndoBERT using Multi-Objective Bayesian Optimization (MOBO) to improve the macro-F1 score and reduce p95 inference latency. The dataset consists of 3,269 labeled positive and negative reviews collected from the Zahira Media Publisher website, which were split using a stratified approach (70% training, 15% validation, and 15% testing). The IndoBERT model (indobenchmark/indobert-base-p1) was trained as a baseline and subsequently optimized using Optuna through the Tree-structured Parzen Estimator (TPE) and Random Sampler within a multi-objective framework. Evaluation was conducted using macro-F1 score and p95 inference latency.

The baseline model achieved a macro-F1 score of 0.8715 with a p95 latency of 25.5060 ms, while the optimized model reached a macro-F1 score of 0.9977 and a p95 latency of 11.8410 ms. These results indicate that MOBO significantly improves classification performance while simultaneously reducing inference time, making the optimized IndoBERT configuration suitable for real-time sentiment analysis systems at Zahira Media Publisher.

Keywords: IndoBERT, Bayesian Optimization, Multi-Objective Optimization, Sentiment Analysis, Optuna