

ABSTRAK

PERBANDINGAN *GRID SEARCH*, *RANDOM SEARCH*, DAN *BAYESIAN OPTIMIZATION* DALAM OPTIMASI *HYPERPARAMETER* MODEL BERT UNTUK ANALISIS SENTIMEN PADA PLATFORM X TERHADAP *WEBSITE CORETAX*

Ady Septian Nugraha
H1D021047

Layanan perpajakan digital *Website Coretax* yang baru disahkan pada awal tahun 2025 memicu beragam opini masyarakat di platform media sosial X. Analisis sentimen terhadap opini tersebut penting dilakukan untuk mengevaluasi kualitas layanan. Penelitian ini bertujuan untuk membandingkan efektivitas tiga metode optimasi *hyperparameter*, yaitu *Grid Search*, *Random Search*, dan *Bayesian Optimization*, pada model *Bidirectional Encoder Representations from Transformers* (BERT) untuk klasifikasi sentimen. Model yang digunakan adalah *indobert-base-p1* dengan *dataset* sebanyak 31.586 *tweet* yang dibagi dengan rasio 80:10:10. Parameter yang dioptimasi meliputi *learning rate*, *batch size*, *epochs*, dan *weight decay*. Hasil penelitian menunjukkan bahwa ketiga metode optimasi berhasil meningkatkan performa dibandingkan *base model* (akurasi 93,84%, presisi 88,98%, *recall* 87,94%, *F1-Score* 88,44%). *Bayesian Optimization* terbukti menjadi metode terbaik dengan peningkatan performa paling signifikan, menghasilkan akurasi 95,00%, presisi 91,86%, *recall* 91,59%, dan *F1-score* 91,72% pada kombinasi *learning rate* $1,04 \times 10^{-5}$, *epoch* 4, *batch size* 16, dan *weight decay* 0,22. Analisis sentimen menunjukkan bahwa mayoritas pengguna memberikan respons negatif terhadap *Website Coretax* sebesar 47% (10.428 *tweet*). Selain itu, penelitian ini mengimplementasikan *dashboard* interaktif berbasis *website* untuk visualisasi hasil analisis dan model.

Kata Kunci: Analisis Sentimen, Coretax, BERT, Optimasi *Hyperparameter*

ABSTRACT

A COMPARISON OF GRID SEARCH, RANDOM SEARCH, AND BAYESIAN OPTIMIZATION IN BERT MODEL HYPERPARAMETER OPTIMIZATION FOR SENTIMENT ANALYSIS ON PLATFORM X TOWARD THE CORETAX WEBSITE

Ady Septian Nugraha
H1D021047

The digital tax service, Coretax Website, launched in early 2025, has triggered various public opinions on the X social media platform. Sentiment analysis of these opinions is crucial for evaluating service quality. This study aims to compare the effectiveness of three hyperparameter optimization methods—Grid Search, Random Search, and Bayesian Optimization—on the Bidirectional Encoder Representations from Transformers (BERT) model for sentiment classification. The model used is indobert-base-p1 with a dataset of 31,586 tweets split at an 80:10:10 ratio. The optimized parameters include learning rate, batch size, epochs, and weight decay. The results show that all three optimization methods improved performance compared to the base model (Accuracy 93.84%, Precision 88.98%, Recall 87.94%, F1-Score 88.44%). Bayesian Optimization proved to be the superior method with the most significant performance increase, achieving 95.00% accuracy, 91.86% precision, 91.59% recall, and a 91.72% F1-score with a configuration of learning rate 1.04×10^{-5} , 4 epochs, batch size 16, and weight decay 0.22. Sentiment analysis revealed that the majority of users expressed negative sentiments toward the Coretax Website at 47% (10,428 tweets). Furthermore, this research implemented an interactive web-based dashboard for visualization of the analysis results and model.

Keywords: *Sentiment Analysis, Coretax, BERT, Hyperparameter Optimization*