

ANALISIS SENTIMEN PENGGUNA *CORETAX* PADA MEDIA SOSIAL MENGGUNAKAN *DECISION TREE* DAN *ADAPTIVE BOOSTING*

Nama Mahasiswa : Nur Safika

Nim 6304221468

Dosen Pembimbing : Ir. Elvi Rahmi, S.T., M.Kom.

ABSTRAK

Implementasi sistem administrasi perpajakan digital terbaru di Indonesia, yaitu *Coretax*, telah memicu beragam opini dan keluhan dari masyarakat di media sosial. Komentar publik yang berjumlah besar dan tidak terstruktur menyulitkan pihak otoritas dalam mengevaluasi layanan secara manual. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap pengguna *Coretax* di platform X (Twitter) dan Instagram guna mengelompokkan opini ke dalam kategori positif, negatif, dan netral. Metode yang digunakan adalah algoritma *Decision Tree* yang dioptimasi dengan teknik ensemble learning Adaptive Boosting (AdaBoost). Tahapan penelitian meliputi web scraping, *preprocessing* teks (termasuk negation merging), pembobotan kata menggunakan TF-IDF, hingga evaluasi model. Dataset yang digunakan berjumlah 1.414 data setelah melalui proses augmentasi untuk menjaga keseimbangan kelas. Hasil penelitian menunjukkan bahwa optimasi AdaBoost berhasil meningkatkan akurasi model sebesar 13,78% dari baseline *Decision Tree*, dengan akurasi akhir mencapai 86,22%. Selain itu, sistem diimplementasikan ke dalam aplikasi berbasis web menggunakan arsitektur hybrid Laravel-Flask yang memungkinkan visualisasi hasil analisis secara interaktif. Penelitian ini diharapkan dapat menjadi alat bantu bagi Direktorat Jenderal Pajak (DJP) dalam memahami pola keluhan pengguna dan meningkatkan kualitas layanan perpajakan digital secara real-time.

Kata kunci: Analisis Sentimen, *Coretax*, *Decision Tree*, Adaptive Boosting, Media Sosial, TF-IDF.

SENTIMENT ANALYSIS OF CORETAX USERS ON SOCIAL MEDIA USING DECISION TREE AND ADAPTIVE BOOSTING

Name : Nur Safika
Student ID : 6304221468
Supervisor : Ir. Elvi Rahmi, S.T., M.Kom

ABSTRACT

The implementation of Indonesia's latest digital tax administration system, Coretax, has sparked various opinions and complaints from the public on social media. The large volume of unstructured public comments makes it difficult for authorities to evaluate services manually. This study aims to perform sentiment analysis on Coretax users on X (Twitter) and Instagram platforms to classify opinions into positive, negative, and neutral categories. The method employed is the Decision Tree algorithm optimized with the Adaptive Boosting (AdaBoost) ensemble learning technique. The research stages include web scraping, text preprocessing (including negation merging), word weighting using TF-IDF, and model evaluation. The dataset consists of 1,414 records after undergoing an augmentation process to maintain class balance. The results show that AdaBoost optimization successfully increased the model's accuracy by 13.78% from the Decision Tree baseline, reaching a final accuracy of 86.22%. Furthermore, the system is implemented into a web-based application using a hybrid Laravel-Flask architecture, allowing for interactive visualization of analysis results. This research is expected to serve as a tool for the Directorate General of Taxes (DJP) to understand user complaint patterns and improve the quality of digital tax services in real-time.

Keywords: Sentiment Analysis, Coretax, Decision Tree, Adaptive Boosting, Social Media, TF-IDF.