

ANALISIS SENTIMEN OPINI PUBLIK TERHADAP INOVASI TEKNOLOGI PADA ADMINISTRASI PUBLIK INDONESIA MENGUNAKAN METODE INDOBERT

Oleh

Komang Tya Vinandita Kusuma Dewi, NIM 2215091044

Jurusan Teknik Informatika

Program Studi Sistem Informasi

ABSTRAK

Penerapan Sistem Pemerintahan Berbasis Elektronik (SPBE) di Indonesia ditujukan untuk meningkatkan efisiensi dan kualitas pelayanan publik. Namun, perbedaan pengalaman pengguna di lapangan mendorong masyarakat untuk menyampaikan aspirasi, kritik, dan keluhan melalui media sosial seperti X dan Tiktok. Penelitian ini bertujuan untuk mengklasifikasi sentimen opini publik terhadap inovasi teknologi administrasi publik di Indonesia menggunakan metode *deep learning* IndoBERT, mengidentifikasi topik-topik dominan menggunakan *Latent Dirichlet Allocation* (LDA), serta mengevaluasi dan mengoptimalkan performa model melalui eksperimen *hyperparameter*, *K-fold*, dan *Earlystopping*. Pengumpulan data dilakukan secara *real-time* dari platform X dan Tiktok yang menghasilkan 11.787 data opini yang setelah melalui tahap seleksi manual dan *preprocessing* menghasilkan 5.541 data bersih dengan distribusi sentimen yaitu sentimen negatif sebanyak 3.566 data dan sentimen positif sebanyak 1.975 data. Ketidakseimbangan kelas sentimen diatasi dengan menggunakan teknik augmentasi data *Back Translation* (BT) menggunakan model *MarianMT*. Hasil penelitian menunjukkan bahwa pengujian pemodelan IndoBERT dengan kombinasi *hyperparameter* 4 *epoch* dan *learning rate* $2e-5$ memberikan performa baseline terbaik dengan akurasi 75% dan F1-Score 74%. Evaluasi *K-fold* menunjukkan kestabilan terbaik pada *epoch* 2 dengan *gap* antara *train accuracy* dan *validation accuracy* terkecil sebesar 12%. Proses optimalisasi model dengan penambahan *epoch* serta penerapan mekanisme *earlystopping* berhasil meningkatkan akurasi model hingga 83% pada *epoch* ke-11. Sementara itu, analisis topik menggunakan LDA berhasil mengidentifikasi 5 topik dominan pada tiap kategori sentimen, dimana sentimen negatif didominasi oleh topik mengenai kendala teknis seperti *error*, *login*, dan verifikasi OTP. Sedangkan pada sentimen positif didominasi oleh apresiasi terhadap kemudahan dan efisiensi layanan digital seperti pembayaran pajak online dan layanan JKN. Seluruh hasil analisis dan evaluasi model telah divisualisasikan ke dalam sebuah *dashboard* berbasis *website*.

Kata kunci: Analisis Sentimen, Administrasi Publik, IndoBERT, LDA, *Dashboard*, *Back Translation*

SENTIMENT ANALYSIS OF PUBLIC OPINION ON TECHNOLOGICAL INNOVATION IN INDONESIAN PUBLIC ADMINISTRATION USING THE INDOBERT METHOD

By

Komang Tya Vinandita Kusuma Dewi, NIM 2215091044

Informatics Engineering

Information System Study Program

ABSTRACT

The implementation of the Electronic-Based Government System (Sistem Pemerintahan Berbasis Elektronik / SPBE) in Indonesia aims to enhance efficiency and public service quality. However, varying user experiences have led the public to express their aspirations, criticisms, and complaints on social media platforms such as X and TikTok. This study aims to classify public opinion sentiment regarding technological innovation in Indonesian public administration using the IndoBERT deep learning method, identify dominant topics using Latent Dirichlet Allocation (LDA), and evaluate as well as optimize model performance through hyperparameter experiments, K-Fold Cross Validation, and Early Stopping mechanisms. Data collection conducted in real-time from X and TikTok gathered 11,787 raw opinions, which yielded 5,541 clean records after manual filtering and preprocessing, consisting of 3,566 negative sentiment data and 1,975 positive sentiment data. Sentiment class imbalance was addressed using data augmentation via Back Translation (BT) with the MarianMT model. The experimental results indicated that the IndoBERT baseline model achieved its best performance with a hyperparameter combination of 4 epochs and a learning rate of $2e-5$, attaining an accuracy of 75% and an F1-score of 74%. Evaluation via K-Fold Cross Validation demonstrated optimal stability at epoch 2, yielding the lowest train-validation accuracy gap of 12%. Model optimization through an increased epoch limit alongside an early stopping mechanism successfully boosted classification accuracy to 83% at the 11th epoch. Furthermore, LDA topic modeling revealed 5 dominant topics per sentiment class, where negative sentiment was predominantly driven by technical issues (such as system errors, login difficulties, and OTP verification), while positive sentiment centered around appreciation for the ease and efficiency of digital services (such as online tax payments and JKN services). All analytical outputs and model evaluation metrics have been visualized within an interactive web-based dashboard.

Keywords: Sentiment Analysis, Public Administration, IndoBERT, Latent Dirichlet Allocation (LDA), Back Translation, SPBE, Dashboard.