

ANALISIS SENTIMEN TERHADAP PERENCANAAN PEMILU ELEKTRONIK DI INDONESIA MENGGUNAKAN PENDEKATAN SOFT VOTING ENSEMBLE

Oleh

I Putu Dennis Prana Arta, NIM 2215091057

Jurusan Teknik Informatika

Program Studi Sistem Informasi

ABSTRAK

Wacana perencanaan Pemilu Elektronik di Indonesia memunculkan respons publik yang beragam di media sosial karena dinilai berpotensi meningkatkan efisiensi penyelenggaraan pemilu, tetapi juga menimbulkan kekhawatiran terkait keamanan data, transparansi, manipulasi, dan kepercayaan terhadap hasil suara. Penelitian ini bertujuan untuk mengidentifikasi sentimen publik, mengevaluasi performa metode *Soft Voting Ensemble*, serta menyajikan hasil analisis melalui topic modeling dan dashboard visualisasi. Penelitian kuantitatif ini menggunakan kerangka kerja CRISP-DM. Data diperoleh dari platform X dan TikTok melalui crawling dan scraping dengan kata kunci terkait pemilu elektronik, e-voting, pemilu digital, dan digitalisasi pemilu. Setelah proses seleksi, diperoleh 3.979 data yang terdiri atas 1.446 sentimen positif dan 2.533 sentimen negatif. Data dilabeli secara manual oleh tiga ahli bahasa, kemudian diproses melalui text preprocessing, pembobotan TF-IDF, pembagian data 80:20 secara stratified, serta pengujian Logistic Regression, *Random Forest*, dan *Soft Voting Ensemble* dengan dan tanpa *Borderline-SMOTE*. Evaluasi dilakukan menggunakan confusion matrix, classification report, dan 10-Fold Cross Validation. Hasil penelitian menunjukkan bahwa *Soft Voting Ensemble* Optimasi Awal dengan *Borderline-SMOTE* memperoleh akurasi dan F1-score tertinggi, yaitu 0,8078 dan 0,8084, tetapi memiliki CV Gap 0,1190 yang mengindikasikan potensi overfitting. Model yang lebih stabil adalah *Soft Voting Ensemble* Tuning tanpa *Borderline-SMOTE* dengan akurasi 0,8003, F1-score 0,7917, dan CV Gap 0,0682. Hasil LDA menunjukkan bahwa sentimen negatif didominasi isu kecurangan, manipulasi data, kebocoran data, keamanan akun, dan kepercayaan terhadap hasil suara, sedangkan sentimen positif berkaitan dengan efisiensi, kemudahan, modernisasi, dan pengurangan penggunaan kertas. Temuan ini menunjukkan bahwa persepsi publik terhadap Pemilu Elektronik masih cenderung skeptis, meskipun terdapat dukungan terhadap efisiensi teknologi. Penelitian selanjutnya disarankan mengoptimalkan model *Soft Voting Ensemble* serta menambahkan evaluasi kualitas topik agar hasil klasifikasi dan interpretasi topik menjadi lebih akurat.

Kata Kunci: Analisis Sentimen, Pemilu Elektronik, Soft Voting Ensemble, Borderline-SMOTE, Topic Modeling.

SENTIMENT ANALYSIS OF ELECTRONIC ELECTION PLANNING IN INDONESIA USING A SOFT VOTING ENSEMBLE APPROACH

By

I Putu Dennis Prana Arta, NIM 2215091057

Department of Informatics Engineering

Information Systems Study Program

ABSTRACT

The discourse on Electronic Election planning in Indonesia has generated diverse public responses on social media because it is considered to have the potential to improve election efficiency, while also raising concerns about data security, transparency, manipulation, and public trust in vote results. This study aims to identify public sentiment, evaluate the performance of the Soft Voting Ensemble method, and present the analysis results through topic modeling and dashboard visualization. This quantitative study applies the CRISP-DM framework. Data were obtained from X and TikTok through crawling and scraping using keywords related to electronic elections, e-voting, digital elections, and election digitalization. After the selection process, 3,979 data points were obtained, consisting of 1,446 positive sentiments and 2,533 negative sentiments. The data were manually labeled by three language experts and then processed through text preprocessing, TF-IDF weighting, an 80:20 stratified data split, and model testing using Logistic Regression, Random Forest, and Soft Voting Ensemble with and without Borderline-SMOTE. Evaluation was conducted using a confusion matrix, classification report, and 10-Fold Cross Validation. The results show that the Initial Optimization Soft Voting Ensemble with Borderline-SMOTE achieved the highest accuracy and F1-score, namely 0.8078 and 0.8084, but produced a CV Gap of 0.1190, indicating potential overfitting. The more stable model was the Tuned Soft Voting Ensemble without Borderline-SMOTE, with an accuracy of 0.8003, an F1-score of 0.7917, and a CV Gap of 0.0682. The LDA results indicate that negative sentiment is dominated by issues of fraud, data manipulation, data leakage, account security, and trust in vote results, whereas positive sentiment is associated with efficiency, convenience, modernization, and reduced paper use. These findings indicate that public perception of Electronic Elections remains generally skeptical, although support for technological efficiency is also present. Future research is recommended to optimize the Soft Voting Ensemble model and incorporate topic quality evaluation to improve classification performance and topic interpretation.

Keywords: Sentiment Analysis, Electronic Election, Soft Voting Ensemble, Borderline-SMOTE, Topic Modeling.