

ABSTRAK

Peran media sosial dalam membentuk opini publik semakin signifikan, khususnya dalam konteks politik lokal seperti Pemilihan Kepala Daerah (Pilkada) Jawa Timur 2024. Penelitian ini bertujuan untuk membangun model klasifikasi sentimen terhadap para kandidat gubernur dan wakil gubernur berdasarkan unggahan di media sosial *X*. Metode yang digunakan adalah algoritma *Random Forest*, yang diaplikasikan melalui tahapan preprocessing teks, pelabelan sentimen, pembobotan fitur menggunakan *TF-IDF*, serta seleksi fitur dengan *Chi-Square* dan *Mutual Information*. Untuk mengatasi ketidakseimbangan kelas, diterapkan teknik sampling seperti *Random Undersampling*, *SMOTE*, *Borderline-SMOTE*, dan kombinasi *SMOTE + Tomek Links*, masing-masing diuji pada rasio data 90:10, 80:20, dan 70:30. Evaluasi performa dilakukan menggunakan confusion matrix dengan metrik akurasi, presisi, *recall*, dan *F1-score*. Hasil terbaik diperoleh pada kombinasi *TF-IDF* dengan *Borderline SMOTE* pada rasio 80:20, dengan akurasi mencapai 82,40% dan *F1-score* 82,44%. Temuan ini menunjukkan bahwa strategi *balancing* yang tepat berperan penting dalam meningkatkan efektivitas model. Selain itu, penelitian ini juga mengimplementasikan dashboard interaktif berbasis Streamlit yang menampilkan visualisasi distribusi sentimen dan prediksi sentimen input secara langsung, sehingga memudahkan eksplorasi data oleh pengguna.

Kata Kunci: Pilkada Jawa Timur, Analisis Sentimen, Media Sosial *X*, *Random Forest*, Seleksi Fitur, Teknik *Balancing*.