

ABSTRAK

Isu lesbian, *gay*, biseksual, dan *transgender* (LGBT) di Indonesia masih tergolong tabu dan sering memicu berbagai reaksi di masyarakat, terutama di media sosial. Fenomena ini mendorong perlunya pemetaan opini publik terhadap isu tersebut melalui analisis sentimen. Penelitian ini bertujuan untuk membandingkan kinerja tiga model *transformer* berbahasa Indonesia yaitu IndoBERT, IndoRoBERTa, dan NusaBERT dalam menganalisis sentimen terhadap isu LGBT pada platform X. Metode yang digunakan mengacu pada pendekatan *Cross Industry Standard Process for Data Mining*. Hasil pengujian dengan menggunakan 6.000 data yang sudah di proses dan dilabeli secara manual maupun otomatis, menunjukkan bahwa model IndoBERT memberikan kinerja terbaik dengan akurasi 89,00% dan *F1-score* sebesar 64,37%. Dan model terbaik ini kemudian digunakan untuk mengklasifikasikan 14.594 *tweet*, dan ditemukan bahwa 91,61% mengandung sentimen negatif, 6,22% netral, dan hanya 2,17% positif. Temuan ini menegaskan bahwa persepsi negatif masyarakat Indonesia terhadap isu LGBT masih sangat dominan di ruang digital.

Kata Kunci: Analisis Sentimen, IndoBERT, IndoRoBERTa, LGBT, NusaBERT



ABSTRACT

The issue of lesbian, gay, bisexual, and transgender (LGBT) in Indonesia remains largely taboo and often triggers various public reactions, especially on social media. This phenomenon underlines the need to map public opinion on the topic through sentiment analysis. This study aims to compare the performance of three Indonesian-language transformer models—IndoBERT, IndoRoBERTa, and NusaBERT—in analyzing sentiment toward LGBT-related discourse on platform X. The research follows the Cross Industry Standard Process for Data Mining (CRISP-DM) approach. The evaluation was conducted using 6,000 data points that had been processed and labeled both manually and automatically. The results show that IndoBERT achieved the best performance, with an accuracy of 89.00% and an F1-score of 64.37%. This top-performing model was then used to classify 14,594 tweets, revealing that 91.61% contained negative sentiment, 6.22% neutral, and only 2.17% positive sentiment. These findings indicate that negative public perceptions of LGBT issues are still highly dominant in Indonesia’s digital landscape.

Keywords: Sentiment Analysis, IndoBERT, IndoRoBERTa, LGBT, NusaBERT

