Sentiment Analysis of Indonesian-English Code-Mixed X Posts on Gender Equality A Hybrid Approach:
Combining Lexicon and Deep Learning Models
Dilya Salsabila, Drs. Medi, M. Kom.
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

SENTIMENT ANALYSIS OF INDONESIAN-ENGLISH CODE-MIXED X POSTS ON GENDER EQUALITY: A HYBRID APPROACH COMBINING LEXICON-BASED AND DEEP LEARNING MODELS

by

DILYA SALSABILA 22/492256/PA/21098

This study explores sentiment analysis on Indonesian-English code-mixed posts from X (formerly Twitter) discussing gender equality, using a hybrid approach that combines lexicon-based auto-labeling with deep learning classification. The research employs the SentIL lexicon to generate sentiment labels through weak supervision after preprocessing. The resulting dataset is used to train and evaluate three deep learning models, specifically BiLSTM, IndoBERT, and IndoBERTweet to assess their performance in classifying sentiment. Preprocessing steps include dropping columns and rows, text cleaning, tokenization, normalization, and selective translation. Model performance is measured using accuracy, precision, recall, F1-score, and Cohen's Kappa, with a manually labeled subset serving as the benchmark. Experimental results show that transformer-based models outperform the BiLSTM baseline, with IndoBERT achieving the highest performance, reaching a macro F1 score of up to 0.68 and an accuracy of 73% uder class weighted training, IndoBERTweet with macro F1 scores up to 0.56, while BiLSTM consistently underperforms, with macro F1 scores remaining below 0.52 across all configurations. These findings indicate that lexicon driven weak supervision is sufficient to train effective sentiment classifiers in low resource, code-mixed settings, while highlighting the advantages of contextual transformer models for capturing nuanced sentiment in gender equality discourse.

Key words: NLP, Sentiment analysis, code-mixed text, gender equality, SentIL, BiLSTM, IndoBERT, IndoBERTweet, Indonesian-English