Abstract

The Ibu Kota Nusantara (IKN) relocation project aims to equalize economic development and reduce the burden on Jakarta, but has elicited mixed reactions from the public, including both support and opposition. Therefore, this study applies machine learning-based sentiment analysis, using Logistic Regression to explore public opinion on the relocation, and leveraging social media data from platform X to gain insights into information, opinions, and public reactions. The Textblob, VADER, and SentiWordNet labeling methods employ a majority vote of the three labels to determine the final label. In order to achieve data balance, SMOTE is employed in this study. Moreover, this study applies a combination of preprocessing, N-gram, and TF-IDF to illuminate the impact of this combination on model performance. The results indicate that the combination of preprocessing Scenario 3 with unigram, bigram, trigram, and TF-IDF feature extraction yields the best performance, achieving a precision of 0.7641, recall of 0.7631, F1-score of 0.7634, and accuracy of 0.7641. This research demonstrates the efficacy of proper preprocessing and feature extraction in enhancing the performance of the Logistic Regression model for sentiment classification, thereby contributing to the analysis of public opinion on IKN policy regarding other issues in the future.

Keywords : Public Opinion, IKN Sentiment, Logistic Regression, Preprocessing, N-gram, TF-IDF