

ABSTRACT

The 2024 General Election in Indonesia sparked various political dynamics widely discussed on platform X. Public responses to the election results reflected satisfaction, hope, and criticism toward the elected president and vice president. These public perceptions serve as important indicators for building trust in the government. Sentiment analysis is essential to help the government monitor public opinion and formulate policies that are both well-targeted and socially acceptable. This study applies Term Frequency–Inverse Document Frequency (TF-IDF) with an N-Gram approach to improve contextual representation in textual data. This approach captures word relationships not recognized by standard TF-IDF. Data were collected using the keywords “Prabowo” and “Gibran” and were classified into three sentiment classes: negative, neutral, and positive. Support Vector Machine (SVM) was chosen for its high classification accuracy. Since SVM is inherently a binary classifier, the One Versus Rest (OVR) strategy was applied to handle the multiclass problem. Four SVM kernels linear, sigmoid, RBF, and polynomial were evaluated. The linear kernel with N-Gram (1,3) was selected as the best-performing model, achieving an accuracy of 76.6%, precision of 79.4%, recall of 76.6%, and F1-score of 74.6%. This outperformed the baseline model using linear kernel and N-Gram (1,1), which achieved 75% accuracy.

Keywords: The 2024 General Election, SVM, OVR, N-Gram, TF-IDF