Analysis of Intolerance to the "Sound Horeg" Phenomenon in Media X Using the SVM and KNN Methods

Dibimbing oleh Muhammad Ainul Fikri, S.T., M.Eng.

Syamaidzar Adani Syah
Program Studi teknik Informatika
Jurusan Teknologi Informasi

ABSTRAK

The "sound horeg" phenomenon on platform X has given rise to a wide range of speech, ranging from rational criticism to intolerant speech that is difficult to distinguish automatically because it is often disguised by subtle linguistic strategies such as sarcasm, idioms, and curses. This study analyzes intolerant discourse in the "sound horeg" phenomenon using Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Ensemble (2:1 soft voting) models, based on 1,393 Indonesian-language tweets labeled using a hybrid method (lexicon and manual) with a Cohen's Kappa of 0.655. Evaluation results on 279 test data points showed that the Ensemble model achieved the highest accuracy at 63.80%, followed by SVM at 63.44% and KNN at 51.25%, with the dominant error pattern being 58.82% of negative statements classified as neutral. An in-depth analysis through the extraction of discriminative keywords validated that the "sound horeg" discourse reflects complex digital social tensions and consists of three interrelated dimensions: the diversity of forms of negative speech, the dimension of physical and health disturbances, and the dimension of conflicts over cultural values and traditions, with the political dimension emerging as an additional finding from the extraction of discriminative keywords. Consequently, the limitations of TF -IDF model in capturing the gray areas of intolerance reflects the complexity of X media users' expressions

Keywords: Sound Horeg, Intolerance Analysis, Support Vector Machine, K-Nearest Neighbor, Ensemble Learning, TF-IDF, Social Media