

ABSTRACT

Academic stress is a common phenomenon among students, characterized by physical and emotional exhaustion. Social media has now become a platform for students to express their academic burdens. This study aims to classify the sentiments of social media users on X related to academic stress using machine learning. Most existing studies have not specifically addressed the use of SVM in relation to academic stress, exploring the influence of stemming and lemmatization, feature extraction techniques, and SVM kernel parameters on sentiment classification performance. This study applies SVM with four types of kernels (linear, RBF, polynomial, and sigmoid). Three experimental stages were conducted: text preprocessing testing, including stemming, lemmatization, and a combination of both; feature extraction technique testing using a combination of TF-IDF with Unigram, Bigram, and Trigram; and SVM kernel and parameter testing, including variations in C, gamma, and degree values. Model performance was measured using accuracy, precision, recall, and F1-score metrics. The best results were obtained from the Polynomial kernel with C=1, gamma=1, and degree=3, achieving an accuracy of 80.79%. On average, the Linear kernel showed the highest stability (65.07%). The combination of stemming + lemmatization preprocessing and Unigram + TF-IDF features proved to be the most effective text representation in academic stress sentiment classification.

Keywords: *academic stress, X, SVM, text preprocessing, N-gram, TF-IDF*