

ABSTRAK

Program Makan Bergizi Gratis (MBG) merupakan salah satu kebijakan pemerintah yang menuai beragam tanggapan dari masyarakat di media sosial X, sehingga diperlukan analisis sentimen untuk mengetahui kecenderungan opini publik terhadap kebijakan tersebut. Penelitian ini bertujuan membandingkan performa dua metode pembobotan fitur, yaitu *Term Frequency-Inverse Document Frequency* (TF-IDF) dan *Word2Vec*, dalam mendukung algoritma *Support Vector Machine* (SVM) untuk mengklasifikasikan sentimen masyarakat menjadi positif dan negatif, menggunakan data sebanyak 10.016 *tweet* yang dikumpulkan melalui teknik *scraping* dengan *web scrapper* Apify dan berkurang menjadi 8.483 data setelah proses *cleaning*, kemudian diproses melalui tahapan *text preprocessing* (*cleaning*, *case folding*, normalisasi, *tokenizing*, *stopword removal*, *stemming*) dan diberi label secara otomatis menggunakan leksikon InSet, dibagi menjadi data latih dan data uji, diseimbangkan dengan *Synthetic Minority Over-sampling Technique* (SMOTE), lalu direpresentasikan menggunakan TF-IDF dan *Word2Vec* berarsitektur *Skip-gram* sebelum diklasifikasikan dengan SVM berkernel *linear*. Hasil penelitian menunjukkan bahwa model SVM dengan pembobotan TF-IDF memberikan performa yang lebih baik dibandingkan *Word2Vec*, dengan nilai *accuracy* sebesar 86%, *precision* sebesar 90%, *recall* sebesar 88%, dan *F1-score* sebesar 89% untuk TF-IDF, sedangkan *Word2Vec* hanya memperoleh *accuracy* sebesar 74%, *precision* sebesar 83%, *recall* sebesar 76%, dan *F1-score* sebesar 79%, sehingga dapat disimpulkan bahwa TF-IDF merupakan metode pembobotan fitur yang lebih efektif dibandingkan *Word2Vec* dalam mendukung algoritma SVM untuk analisis sentimen terhadap Program Makan Bergizi Gratis pada data media sosial X.

Kata kunci: Analisis Sentimen, TF-IDF, *Word2Vec*, *Support Vector Machine*, Program Makan Bergizi Gratis, Media Sosial X

ABSTRACT

The Free Nutritious Meal Program (*Makan Bergizi Gratis* / MBG) is a government policy that has generated diverse public responses on social media platform X, making sentiment analysis necessary to determine the trend of public opinion toward the policy. This study aims to compare the performance of two feature weighting methods, *Term Frequency-Inverse Document Frequency* (TF-IDF) and *Word2Vec*, in supporting the *Support Vector Machine* (SVM) algorithm to classify public sentiment into positive and negative categories, using 10,016 *tweets* collected through a *scraping* technique with the *Apify web scraper* and reduced to 8.483 data after the *cleaning* process, which were then processed through *text preprocessing* stages (*cleaning*, *case folding*, *normalization*, *tokenizing*, *stopword removal*, *stemming*) and automatically labeled using the InSet lexicon, split into training and testing sets, balanced using the *Synthetic Minority Over-sampling Technique* (SMOTE), and represented using TF-IDF and *Word2Vec* with a *Skip-gram* architecture before being classified using SVM with a *linear kernel*. The results show that the SVM model with TF-IDF weighting performed better than *Word2Vec*, achieving an *accuracy* of 86%, *precision* of 90%, *recall* of 88%, and *F1-score* of 89% for TF-IDF, while *Word2Vec* only achieved an *accuracy* of 74%, *precision* of 83%, *recall* of 76%, and *F1-score* of 79%, so it can be concluded that TF-IDF is a more effective feature weighting method than *Word2Vec* in supporting the SVM algorithm for sentiment analysis of the Free Nutritious Meal Program on social media X data.

Keywords: Sentiment Analysis, TF-IDF, Word2Vec, Support Vector Machine, Free Nutritious Meal Program, Social Media X