

ABSTRAK

Perhatian publik terhadap pelaksanaan Program Makan Bergizi Gratis (MBG) meningkat seiring luasnya cakupan program dan dampaknya *terhadap* masyarakat. Penelitian ini bertujuan untuk menganalisis sentimen publik secara multi-aspek terhadap Program MBG menggunakan algoritma *Logistic Regression* dengan representasi fitur *TF-IDF*, pendekatan *N-Gram*, serta penerapan teknik *SMOTE* untuk mengatasi ketidakseimbangan kelas sentimen. Data penelitian diperoleh dari komentar publik di media sosial TikTok sebanyak 1.390 data yang telah melalui proses pelabelan ke dalam lima aspek utama, yaitu Kualitas dan Keamanan Menu, Anggaran dan Efisiensi, Ketepatan Waktu Distribusi, Kesiapan Infrastruktur dan Logistik, serta Cakupan dan Pemerataan. Ekstraksi fitur dilakukan menggunakan *TF-IDF* dengan variasi konfigurasi *N-Gram*, dan klasifikasi sentimen dilakukan menggunakan *Logistic Regression*. Hasil pengujian menunjukkan bahwa penggunaan kombinasi unigram–bigram (*N-Gram range (1,2)*) menghasilkan performa terbaik dengan rata-rata akurasi sebesar 86,1%, rata-rata presisi 85,0%, rata-rata recall 81,6%, dan rata-rata F1-score 82,8% pada data uji. Penerapan metode *SMOTE* terbukti meningkatkan kinerja model secara signifikan, khususnya pada metrik recall dan F1-score kelas minoritas (positif). Nilai rata-rata F1-score meningkat dari 69,3% pada model tanpa *SMOTE* menjadi 82,8% setelah penerapan *SMOTE*, atau mengalami peningkatan sebesar 13,5%. Temuan ini menunjukkan bahwa kombinasi *TF-IDF*, *N-Gram*, dan *SMOTE* dengan algoritma *Logistic Regression* efektif dalam melakukan analisis sentimen multi-aspek terhadap Program MBG dan dapat menjadi dasar evaluasi berbasis data bagi pemangku kebijakan.

Kata kunci: *Analisis Sentimen, Multi-Aspek, MBG, TF-IDF, N-Gram, Logistic Regression, Klasifikasi Teks*

ABSTRACT

Public attention to the implementation of the Free Nutritious Meal Program (Program Makan Bergizi Gratis/MBG) has increased along with the program's wide coverage and its impact on society. This study aims to analyze public sentiment toward the MBG program using a multi-aspect approach by applying Logistic Regression with TF-IDF feature representation, an N-Gram approach, and the Synthetic Minority Over-sampling Technique (SMOTE) to address sentiment class imbalance. The dataset consists of 1,390 public comments collected from the TikTok social media platform, which were manually labeled into five main aspects: Menu Quality and Safety, Budget and Efficiency, Distribution Timeliness, Infrastructure and Logistics Readiness, and Coverage and Equity. Feature extraction was performed using TF-IDF with various N-Gram configurations, and sentiment classification was conducted using the Logistic Regression algorithm. Experimental results show that the combination of unigram–bigram features (N-Gram range (1,2)) achieved the best performance, with an average accuracy of 86.1%, average precision of 85.0%, average recall of 81.6%, and an average F1-score of 82.8% on the test data. The application of SMOTE significantly improved model performance, particularly in recall and F1-score for the minority (positive) class, with an average F1-score improvement of 13.5% compared to the model without SMOTE. These findings indicate that the integration of TF-IDF, N-Gram, and SMOTE with Logistic Regression is effective for multi-aspect sentiment analysis of the MBG program and can serve as a data-driven evaluation basis for policymakers.

Keywords: *Sentiment Analysis, Multi-Aspect, MBG, TF-IDF, N-Gram, Logistic Regression, Text Classification*