

ABSTRAK

Perkembangan industri kuliner modern sangat dipengaruhi oleh ulasan pelanggan di platform digital, termasuk Google Review. Ulasan ini mencerminkan persepsi konsumen terhadap produk dan layanan yang diberikan oleh suatu usaha, seperti Kaktus Coffee Place. Analisis sentimen menjadi pendekatan yang banyak digunakan untuk memahami opini pelanggan tersebut. Namun, dalam praktiknya, analisis sentimen menghadapi tantangan berupa ketidakseimbangan data, di mana ulasan positif jauh lebih dominan dibandingkan ulasan negatif. Ketidakseimbangan ini dapat menurunkan kinerja model analisis yang digunakan.

Penelitian ini menggunakan algoritma *Support Vector Machine* (SVM) untuk mengklasifikasikan sentimen ulasan Google Kaktus Coffee Place ke dalam kategori positif dan negatif. Untuk mengatasi ketidakseimbangan data, diterapkan teknik *Synthetic Minority Oversampling Technique* (SMOTE) yang menghasilkan data sintetis pada kelas minoritas. Data ulasan dikumpulkan melalui web scraping dengan bantuan platform Apify, kemudian dilakukan preprocessing teks seperti case folding, cleansing, tokenisasi, normalisasi kata, stopword removal, dan stemming. Representasi fitur dilakukan menggunakan TF-IDF, dengan penanganan negasi menggunakan *bigram*. Model kemudian diuji menggunakan metode *K-Fold Cross Validation*.

Hasil pengujian menunjukkan bahwa penerapan SMOTE meningkatkan performa model analisis sentimen secara signifikan dibandingkan model tanpa penyeimbangan data. Peningkatan terlihat pada metrik evaluasi seperti akurasi, *precision*, *recall*, dan *F1-score*. Penelitian ini memberikan kontribusi nyata dalam bidang analisis sentimen berbasis teks dengan pendekatan pembelajaran mesin, khususnya dalam penanganan data tidak seimbang pada ulasan daring.

Kata Kunci: Analisis Sentimen, SVM, SMOTE, TF-IDF, *Google Review*

ABSTRACT

The modern culinary industry is significantly influenced by customer reviews on digital platforms, particularly Google Reviews. These reviews reflect customer perceptions of the products and services offered by a business, such as Kaktus Coffee Place. Sentiment analysis has become a widely used approach to understanding customer opinions. However, in practice, sentiment analysis often faces the challenge of imbalanced data, where positive reviews significantly outnumber negative ones. This imbalance can reduce the performance of classification models.

This study utilizes the Support Vector Machine (SVM) algorithm to classify sentiment in Google reviews of Kaktus Coffee Place into positive and negative categories. To address the data imbalance, the Synthetic Minority Oversampling Technique (SMOTE) is applied to generate synthetic data in the minority class. The review data was collected through web scraping using the Apify platform, followed by several preprocessing steps, including case folding, cleansing, tokenization, word normalization, stopword removal, and stemming. Feature representation was conducted using TF-IDF, with negation handling implemented through bigrams. The model was evaluated using the K-Fold Cross Validation method.

The evaluation results show that the application of SMOTE significantly improves the performance of the sentiment classification model compared to models trained on imbalanced data. Improvements were observed in evaluation metrics such as accuracy, precision, recall, and F1-score. This research provides a valuable contribution to text-based sentiment analysis using machine learning approaches, particularly in addressing imbalanced datasets in online reviews.

Keywords: Sentiment Analysis, SVM, SMOTE, TF-IDF, Google Review