

ABSTRAK

Penelitian ini menganalisis sentimen ulasan pelanggan terhadap Rumah Makan Bebek Sinjay di Madura menggunakan metode *Random Forest* dengan teknik *balancing data* SMOTE untuk mengatasi ketidakseimbangan kelas ulasan positif dan negatif. Data ulasan diambil secara otomatis dari *Google Maps* menggunakan teknik *web scraping*, kemudian dilakukan preprocessing meliputi *case folding*, *cleaning*, *tokenizing*, *stopword removal*, dan *stemming*. Selanjutnya, data diubah menjadi representasi *numerik* menggunakan TF-IDF dan dibagi menjadi data latih dan data uji dengan proporsi 80:20. Model *Random Forest* dibangun dan dioptimasi dengan *Grid Search* untuk menemukan *hyperparameter* terbaik. Evaluasi model dengan *confusion matrix* menghasilkan akurasi terbaik sebesar 0.82, presisi 0.82, recall 0.82. Hasil penelitian menunjukkan bahwa penerapan SMOTE meningkatkan performa model dalam mengklasifikasikan sentimen ulasan, terutama dalam mendeteksi ulasan negatif yang jumlahnya lebih sedikit. Temuan ini memberikan wawasan berharga bagi pengelola Bebek Sinjay untuk mengidentifikasi kritik pelanggan secara lebih akurat dan meningkatkan kualitas pelayanan. Penelitian ini menegaskan efektivitas *Random Forest* dalam analisis sentimen pada data teks ulasan berbasis *machine learning*.

Kata kunci : Analisis Sentimen, *Random Forest*, SMOTE, *TF-IDF*, Ulasan Pelanggan, *Machine Learning*.

Abstract

This study analyzes customer review sentiments for Bebek Sinjay Restaurant in Madura using the Random Forest method with the SMOTE data balancing technique to address the imbalance of positive and negative review classes. Review data is automatically retrieved from Google Maps using web scraping techniques, then preprocessed including case folding, cleaning, tokenizing, stopword removal, and stemming. Next, the data is converted into numeric representation using TF-IDF and divided into training data and test data with a proportion of 80:20. A Random Forest model is built and optimized with Grid Search to find the best hyperparameters. Model evaluation with a confusion matrix produces the best accuracy of 0.82, precision 0.82, recall 0.82. The results show that the application of SMOTE improves the model's performance in classifying review sentiments, especially in detecting fewer negative reviews. These findings provide valuable insights for Bebek Sinjay managers to identify customer criticism more accurately and improve service quality. This study confirms the effectiveness of Random Forest in sentiment analysis on machine learning-based review text data.

Keywords: *Sentiment Analysis, Random Forest, SMOTE, TF-IDF, Customer Reviews, Machine Learning.*