

## ABSTRACT

Customer reviews are opinions on the quality of goods or services obtained by consumers. Customer reviews contain useful information for consumers and providers of goods or services. The availability of a large number of customer reviews on online media requires a framework to extract sentiments automatically. A customer review often contains many aspects so that Aspect Based Sentiment Analysis (ABSA) can be used to determine the polarity of each aspect and sentiment whether it is positive, negative or neutral. The main tasks in aspect-based sentiment analysis are aspect classification and sentiment classification.

Machine learning methods for classifying aspects and sentiments have been widely used in English datasets, but only in Indonesian datasets. This study compares the performance of five machine learning algorithms, namely Support Vector Machine (SVM), Naïve Bayes (NB), Logistic Regression (LR), Random Forest (RF) and Extreme Gradient Boosting (XGBoost) on Indonesian-language customer reviews (restaurant and restaurant domains). e-commerce) using Term Frequency–Inverse Document Frequency (TF-IDF) as term weighting.

This study resulted in the RF method having the most superior performance for aspect classification and sentiment classification compared to the SVM, NB, LR and XGBoost methods in the two domains of restaurant and e-commerce. Aspect classification using the RF method produces an f1-score of 0.879 for the restaurant domain and 0.835 for the e-commerce domain. Sentiment classification using the RF method produces an f1-score of 0.695 for the restaurant domain and 0.680 for the e-commerce domain.

**Keywords:** *Aspect Based Sentiment Anlaysis, Customers Review, Aspect Classification, Sentiment Classification, Machine Learning*

## INTISARI

Ulasan pelanggan merupakan opini terhadap kualitas barang atau jasa yang didapatkan oleh konsumen. Ulasan pelanggan mengandung informasi yang berguna bagi konsumen maupun penyedia barang atau jasa. Ketersediaan ulasan pelanggan dalam jumlah besar pada media *online* membutuhkan suatu *framework* untuk mengekstraksi sentimen secara otomatis. Sebuah ulasan pelanggan sering kali mengandung banyak aspek sehingga *Aspect Based Sentiment Analysis* (ABSA) dapat digunakan untuk mengetahui polaritas masing-masing aspek dan sentimen baik itu positif, negatif ataupun netral. Tugas utama dalam analisis sentimen berbasis aspek adalah klasifikasi aspek dan klasifikasi sentimen.

Metode *machine learning* untuk klasifikasi aspek dan sentimen sudah banyak dilakukan pada *dataset* berbahasa Inggris, tetapi pada *dataset* berbahasa Indonesia masih sedikit. Penelitian ini membandingkan kinerja lima algoritma *machine learning*, yaitu *Support Vector Machine* (SVM), *Naïve Bayes* (NB), *Logistic Regression* (LR), *Random Forest* (RF) dan *Extreme Gradient Boosting* (XGBoost) pada ulasan pelanggan berbahasa Indonesia (domain restoran dan *e-commerce*) dengan menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) sebagai *term weighting*.

Penelitian ini menghasilkan metode RF memiliki kinerja paling unggul untuk klasifikasi aspek dan klasifikasi sentimen dibandingkan dengan metode SVM, NB, LR dan XGBoost pada dua domain restoran dan *e-commerce*. Klasifikasi aspek dengan metode RF menghasilkan nilai *f1-score* 0,879 untuk domain restoran dan 0,835 untuk domain *e-commerce*. Klasifikasi sentimen dengan metode RF menghasilkan nilai *f1-score* 0,695 untuk domain restoran dan 0,680 untuk domain *e-commerce*.

**Kata kunci** – *Sentimen Analisis Berbasis Aspek, Ulasan Pelanggan, Klasifikasi Aspek, Klasifikasi Sentimen, machine learning.*