

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap produk *Hada Labo Gokujyun Ultimate Moisturizing Face Wash* yang diperoleh dari *platform Female Daily*. Metode yang digunakan adalah algoritma Naive Bayes dengan pendekatan *Natural language Processing* (NLP). Proses NLP mencakup tahap pra-pemrosesan teks menggunakan pustaka Sastrawi, pelabelan otomatis sentimen dengan model IndoBERT, serta ekstraksi fitur menggunakan metode TF-IDF. Dataset penelitian terdiri dari 1.770 ulasan yang dikumpulkan pada periode September 2015 hingga Januari 2025.

Hasil pengujian menunjukkan bahwa algoritma Naive Bayes mampu mengklasifikasikan ulasan ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Model memperoleh akurasi sebesar 97%, presisi 81%, *recall* 83%, dan F1-score 79% (*macro average*). Analisis distribusi sentimen memperlihatkan bahwa mayoritas ulasan bersifat positif (1.611 ulasan), diikuti oleh ulasan negatif (147 ulasan), dan sebagian kecil ulasan netral (12 ulasan).

Sebagai implementasi, penelitian ini menghasilkan aplikasi berbasis web menggunakan Streamlit yang mampu memfasilitasi proses unggah data, pra-pemrosesan, pelabelan sentimen, pelatihan model, evaluasi performa, hingga visualisasi hasil klasifikasi secara interaktif.

Dengan demikian, penelitian ini membuktikan bahwa algoritma Naive Bayes berbasis NLP dapat digunakan secara efektif untuk menganalisis sentimen ulasan konsumen *skincare*, serta memberikan kontribusi pada pengembangan metode analisis teks berbahasa Indonesia di bidang *Natural language Processing*.

Kata Kunci: Analisis Sentimen, Naive Bayes, IndoBERT, *Skincare*, *Female Daily*, Streamlit, NLP

ABSTRACT

This study aims to analyze the sentiment of user reviews on the Hada Labo Gokujyun Ultimate Moisturizing Face Wash product obtained from the Female Daily platform. The method employed is the Naive Bayes algorithm with a Natural language Processing (NLP) approach. The NLP process includes text preprocessing using the Sastrawi library, automatic sentiment labeling with the IndoBERT model, and feature extraction using the TF-IDF method. The research dataset consists of 1,770 reviews collected from September 2015 to January 2025.

The experimental results show that the Naive Bayes algorithm is capable of classifying reviews into three sentiment categories: positive, neutral, and negative. The model achieved an accuracy of 97%, precision of 81%, recall of 83%, and an F1-score of 79% (macro average). Sentiment distribution analysis indicates that the majority of reviews are positive (1,611 reviews), followed by negative (147 reviews), and a small portion are neutral (12 reviews).

As an implementation, this research developed a web-based application using Streamlit that facilitates data uploading, preprocessing, sentiment labeling, model training, performance evaluation, and interactive visualization of classification results.

In conclusion, this study demonstrates that the Naive Bayes algorithm based on NLP can be effectively used to analyze consumer skincare product reviews, while also contributing to the development of text analysis methods for the Indonesian language in the field of Natural language Processing.

Keywords: Sentiment Analysis, Naive Bayes, NLP, IndoBERT, TF-IDF, Skincare, Streamlit