

## ABSTRAK

Produk kecantikan telah menjadi salah satu hal yang tidak bisa lepas dari kaum wanita bahkan menjadi suatu kebutuhan dasar untuk merawat, mempercantik dan menunjang penampilan seseorang. Semakin banyaknya platform *e-commerce* yang menyediakan ulasan atau *review* produk semakin meningkat pula selektivitas konsumen dalam memilih produk. Ulasan suatu produk kecantikan dapat membantu konsumen untuk mengetahui kualitas produk yang ingin digunakan. Namun, dengan banyaknya jumlah ulasan yang tersedia tidak memungkinkan untuk membaca detail satu persatu *review* karena dapat memakan waktu.

Penelitian ini dapat membantu konsumen untuk mengetahui makna sentimen dari ulasan produk dengan melakukan analisis sentimen. Analisis sentimen terdiri dari dua kelas yaitu kelas positif dan kelas negatif. Tahapan analisis sentimen pada penelitian ini dimulai dengan pengumpulan data ulasan yang dilakukan secara manual dari website *sociolla*, kemudian pelabelan data ulasan berdasarkan rating menjadi dua kelas yaitu kelas positif dan kelas negatif, rating 1-2 untuk kelas negatif dan rating 4-5 untuk kelas positif. Selanjutnya data akan dilakukan *preprocessing*, kemudian tahap ekstraksi fitur dengan TF-IDF, kemudian pemodelan dengan metode *Naïve Bayes Classifier*. Model yang sudah dibangun akan dilakukan pengujian terhadap data test dengan tabel *confusion matrix*. Dari tabel *confusion matrix* akan diperoleh nilai akurasi, *presisi* dan *recall*.

Hasil dari pengujian yang menggunakan 200 data ulasan dengan perbandingan data training dan data uji sebesar 90% : 10 % dimana untuk data training 180 data dan data testing 20 data, hasil pengujian dari 10 data ulasan negatif terklasifikasi 9 data dinyatakan negatif sesuai prediksi. Dan 1 data diprediksi negatif tetapi hasilnya positif, begitu dengan 10 data positif terklasifikasi 7 sesuai prediksi dan 3 data diprediksi positif tetapi hasilnya negatif. Dengan *confusion matrix* mendapatkan hasil akurasi yang cukup baik yaitu sebesar 80%, *presisi* 80% dan *recall* 81,25%.

**Kata kunci:** Analisis Sentimen, *Review*, TF-IDF, Naïve Bayes , *Confusion Matrix*

## ABSTRACT

Beauty products have become one of the things that cannot be separated from women, and have even become a basic need to care for, beautify and support one's appearance. The increasing number of e-commerce platforms that provide product reviews or reviews also increases the selectivity of consumers in choosing products. Reviews of a beauty product can help consumers to know the quality of the product they want to use. However, with the large number of available reviews it is not possible to read the details one by one because it can take time.

This research can help consumers to find out the meaning of sentiment from product reviews by conducting sentiment analysis. Sentiment analysis consists of two classes, namely the positive class and the negative class. The stages of sentiment analysis in this study began with collecting review data manually from the sociolla website, then labeling review data based on ratings into two classes, namely positive class and negative class, rating 1-2 for negative class and rating 4-5 for positive class. . Furthermore, the data will be preprocessed, then the feature extraction stage with TF-IDF, then modeling with the Naïve Bayes Classifier method. The model that has been built will be tested on test data with a confusion matrix table. From the confusion matrix table, accuracy, precision and recall values will be obtained.

The results of the test using 200 review data with a comparison of training data and test data of 90%: 10% where for 180 data training data and 20 data testing data, the test results from 10 negative review data classified as 9 data were declared negative according to predictions. And 1 data is predicted to be negative but the result is positive, so with 10 positive data it is classified 7 as predicted and 3 data is predicted to be positive but the result is negative. With the confusion matrix, the accuracy results are quite good, namely 80%, 80% precision and 81.25% recall.

**Keywords:** Sentiment Analysis, Review, TF-IDF, Naïve Bayes, Confusion Matrix