

## ABSTRACT

From 2022 to mid-2023, the beauty industry's development has grown by 21.9%. In line with the increase in the beauty industry, there are many forums to discuss and review these beauty products. One of the forums is Female Daily which reviews can convince consumers to buy these products. Based on Brightlocal data in consumer reviews, as many as 82% of respondents will delay buying a product if there is a negative review. This indicates that the reviews significantly affect product quality. One way to conclude from reviews is with aspect-based sentiment analysis. Therefore, this research performs aspect extraction with token-level classification on beauty product reviews using pre-trained language models, namely IndoBERT and mBERT. The best performances are achieved by IndoBERT with a 69% F1-score, while mBERT has 68% F1-score. IndoBERT is better at handling Indonesian contexts due to its training on a large and specific Indonesian dataset, while mBERT is more flexible for multilingual data but less optimized for Indonesian-specific tasks. In addition, the highest performance for each aspect is obtained by the aspect price of 75% F1-score at B-Price and 71% F1-score at I-Price, compared to the aspects of packaging and texture. This is because the price aspect is easier to recognize as it tends to have more specific terminology, such as cheap, expensive, etc.

Keywords: Aspect Extraction, Beauty Product Reviews, IndoBERT, mBERT, Pre-trained Language Model