Abstract—This paper presents an integrated approach to sentiment analysis of beauty product reviews using the IndoBERT model combined with Naive Bayes classification, which specifically addresses the challenge of accurately analyzing mixed-language reviews. This research outlines the problem of mixed language sentiment analysis and introduces a methodology that includes data preprocessing, feature extraction with IndoBERT, and sentiment classification using Naive Bayes. The results, presented later in this paper, demonstrate that the Gaussian Naive Bayes classifier, in combination with IndoBERT, achieves a Mean Macro F1 Score of 0.569 (56.9%) in the Price category. This integrated approach significantly improves computational efficiency and accuracy, offering valuable insights for natural language processing (NLP) and digital marketing applications.

Index Terms—sentiment analysis, beauty product reviews, IndoBERT, naive bayes, NLP