Abstract−The hospitality industry heavily relies on online reviews as a crucial source of information that influences potential guests' decisions. However, conducting sentiment analysis on hotel reviews can be challenging due to the complexity of language and contextual diversity, especially in Indonesian. This study aims to develop and optimize a RoBERTa-based sentiment analysis model to improve the accuracy of sentiment classification in Indonesian hotel reviews, focusing on the aspects of facilities, cleanliness, location, price, and service. The methodology includes data collection through web scraping from the Traveloka platform, manual labeling, and text pre-processing. The RoBERTa model was trained and optimized using fine-tuning techniques and evaluated using metrics such as accuracy, precision, recall, F1-score, and AUC. The results show that the optimized RoBERTa model achieves competitive performance, although the IndoBERT model with Bayesian Optimization demonstrates superior performance, particularly in terms of accuracy and efficiency in identifying positive and negative sentiments. This study is expected to contribute to the development of more effective and accurate aspect-based sentiment analysis (ABSA) for Indonesian-language hotel reviews. It also opens opportunities for applying NLP technology in the hospitality industry and across other review platforms, thereby improving sentiment analysis quality and assisting hotel managers in enhancing service and customer experience.
Keywords: Aspect-Based Sentiment Analysis; RoBERTa; IndoBERT; Fine-Tuning