ABSTRACT

The growth in railway passenger numbers in Bandung has shown a significant increase, particularly during holidays and weekends. This surge demands improved service quality, facilities, and station infrastructure to maintain user satisfaction and comfort. However, user complaints still arise regarding inadequate amenities, slow service, and poor station cleanliness. Considering the importance of train stations as the primary gateway to public transportation, a systematic and accurate evaluation of public opinion is crucial for PT KAI to make decisions that are more targeted and aligned with actual field needs.

This study proposes a sentiment analysis approach based on user reviews from Google Maps to better understand public perceptions of station services, facilities, and infrastructure in Bandung. Data was collected through a web scraping process using SerpApi. Once gathered, the text data underwent preprocessing to remove irrelevant elements such as numbers, punctuation marks, and stopwords.The next stage involved automatic labeling using the Indonesian-language BERT model, a Transformer-based architecture capable of bidirectional sentence understanding. Each review was categorized into one of three sentiment classes: positive, neutral, or negative. The labeled reviews were then transformed into numerical vectors using TF-IDF (Term Frequency–Inverse Document Frequency). These vectors were used for classification using three machine learning algorithms: Naive Bayes, Support Vector Machine (SVM), and Random Forest. To address class imbalance in the training data, the SMOTE method was applied.

Model performance was evaluated using accuracy, precision, recall, and F1-score metrics, with the best results achieved by the SVM model: 81% accuracy, 73% precision, 71% recall, and 71% F1-score. According to the analysis, the topic "Comfortable waiting room" was the most frequently discussed in positive sentiment reviews (645 mentions), while "User-unfriendly accessibility" was the most discussed in negative sentiment reviews (117 mentions). These findings are expected to serve as strategic input for station management in improving the quality of railway services in the future.

Keywords: Sentiment Analysis, Google Maps, Railway Station, Naive Bayes, Support Vector Machine, Random Forest, BERT, Bandung City