SENTIMENT ANALYSIS OF THE TJ: TRANSJAKARTA APPLICATION USING THE INDOBERT MODEL

Agry Alfiah
Universitas Gunadarma, Indonesia
agry_alfiah@staff.gunadarma.ac.id

Abstract

This research focuses on identifying user sentiment toward the TJ: Transjakarta application based on reviews available on the Google Play Store. A total of 3,431 user reviews were collected using web scraping techniques. The collected data were processed through several preprocessing stages, including text cleaning, normalization, tokenization, stopword removal, and stemming. Sentiment labels were initially assigned using a dictionary-based lexicon approach with the InSet lexicon, resulting in 1,991 positive and 1,440 negative reviews. The dataset was subsequently divided into training, validation, and testing subsets before fine-tuning the IndoBERT-base-p1 model. Experimental results show that the proposed model achieved an accuracy of 87.76%, with balanced precision and recall across both sentiment classes. Further analysis using word cloud visualization indicates that positive sentiment is mainly related to route accessibility and schedule information, while negative sentiment is dominated by technical issues and transaction-related problems. These results confirm the effectiveness of IndoBERT for Indonesian sentiment analysis and its potential use in evaluating the quality of public transportation applications

Keywords: *Sentiment analysis; IndoBERT; TJ: Transjakarta; InSet.*

INTRODUCTION

Public transportation plays a strategic role in supporting the mobility of residents of Jakarta, known as a highly active metropolitan area. The imbalance between the growth of private vehicles and the availability of road infrastructure has led to increasingly complex congestion problems, necessitating improvements in the quality of public transportation services as a sustainable solution to increase travel efficiency and comfort.

As part of its digital transformation efforts for public transportation services, PT Transportasi Jakarta launched the TJ: Transjakarta app in September 2024. This app is designed to help users access various Transjakarta service information, such as real-time bus routes and schedules, digital ticket purchases, and other supporting features. The TJ: Transjakarta app is expected to improve service quality and provide a better experience for public transportation users in Jakarta.

To objectively assess app quality, user reviews available on the Google Play Store can be utilized as a data source because they represent users' experiences and perceptions

of the service provided. These reviews include a variety of responses, both positive and negative, regarding the app's functionality and performance. The large number of reviews makes manual analysis less effective, necessitating an automated approach to systematically categorize user sentiment.

Sentiment analysis is one application of Natural Language Processing (NLP) that aims to identify opinions or emotions contained in text, both positive and negative. In this study, sentiment analysis was conducted using the IndoBERT model, a transformer architecture-based language model specifically designed to more accurately understand the Indonesian language context. The use of IndoBERT is expected to produce better sentiment classification performance than conventional methods, particularly in capturing the meaning and nuances of language.

Based on this background, this study aims to analyze user sentiment toward the TJ: Transjakarta application based on reviews provided on the Google Play Store using the IndoBERT model. The results of this study are expected to provide an overview of user satisfaction levels and serve as evaluation material for application managers in improving service quality and user experience in the future.

METHODS

This study applies a supervised machine learning approach combined with natural language processing to classify sentiment in user reviews of the TJ: Transjakarta app. The model used in this study is IndoBERT-base-p1, adjusted through a fine-tuning process. The research stages include data collection, text preprocessing, sentiment labeling, dataset sharing, model deployment, and performance evaluation.

A. Data Collection

The researchers collected user review data using web scraping techniques on reviews of the TJ: Transjakarta app available on the Google Play Store. The scraping process was carried out using the Python library google_play_scraper. The collected data included review text, user name, rating, and review date. All obtained data was then stored and used as a research dataset.

B. Data Preprocessing

The data preprocessing stage aims to prepare the review text for optimal processing by the model. This process includes cleaning the text by removing links, emojis, punctuation, numbers, and special characters, and converting all text to lowercase. Next, non-standard words are normalized to standard words, stopwords are removed, and stemming is performed to return words to their basic form. The text is then tokenized using the IndoBERT-base-p1 model's tokenizer to produce token representations that align with the model's architecture.

C. Sentiment Labeling

Sentiment labeling was performed using the InSet (Indonesian Sentiment Lexicon) dictionary-based lexicon method. Each token in the review text was matched with words in the InSet dictionary to obtain a polarity value. The polarity values of all tokens were then summed to determine the sentiment label. A review was classified as positive sentiment if the sum of the polarity values was greater than or equal to zero, while reviews with a sum of polarity values less than zero were classified as negative sentiment.

D. Data Division

The labeled dataset was then divided into three parts: training data, validation data, and test data. The data division was carried out with a ratio of 80% for training data, 10% for validation data, and 10% for test data. This division aimed to ensure optimal model training and to allow for objective evaluation of model performance on previously unseen data.

E. IndoBERT Model Implementation

The model implementation was carried out through a fine-tuning process using IndoBERT-base-p1, which is based on the BERT-base architecture. The model was accessed through the Transformers library from HuggingFace, which is compatible with the PyTorch and TensorFlow frameworks. Preprocessed text was converted into a numeric representation using the IndoBERT tokenizer. This representation was then processed in the embedding stage to capture the semantic meaning and context of words within a sentence. The fine-tuning process was carried out by adjusting the model weights to the research dataset.

F. Model Evaluation

Model performance was evaluated using classification metrics available in the Scikit-learn library, specifically through the classification_report function. The metrics used included accuracy, precision, recall, and F1-score. In addition to the quantitative evaluation, word cloud visualization was used to illustrate the dominant words in each sentiment class, while bar charts were used to show the distribution of data within each sentiment class.

RESULTS AND DISCUSSIONS

The data used in this study are user reviews of the TJ: Transjakarta application obtained from the Google Play Store using web scraping techniques. The data collection process resulted in 3,431 reviews collected between September 2024 and October 2025. This data is used as the basis for analyzing user sentiment towards the service quality of the TJ: Transjakarta application.

Figure 1. Results of Data Collection in the TJ Application: Transjakarta

Before sentiment analysis is performed, the review data undergoes a cleaning and preprocessing process to ensure data quality. This process includes removing blanks and duplicates, as well as removing irrelevant characters from the text. The preprocessing process aims to produce cleaner text ready for use in sentiment labeling and model training.

Table 1
Example of Cleaning Text Results

Before Cleaning Text	After Cleaning Text
The app is great; it can see the location of a TJ or JAK and also estimate their arrival time. The downside is that when you use your phone's QR code to scan, it takes a long time to work when tapping in or out. Please fix this, thank you.	The application is good, it can see the location point of TJ or JAK and can also estimate when it will arrive. The drawback is that if you use a cellphone QR code when you want to scan it, it takes a long time to function when tapping in or tapping out. Please fix it, thank you.

Initial sentiment labeling was performed using the InSet dictionary-based lexicon method. The labeling results showed a relatively balanced sentiment distribution, with 1,991 reviews labeled positive and 1,440 reviews labeled negative. This distribution indicates that user perceptions of the TJ: Transjakarta application are not homogeneous, but rather reflect a combination of user satisfaction and complaints.
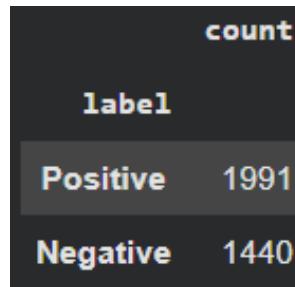
Figure 2. Number of Sentiment Labels

The labeled dataset was then divided into training data, validation data, and test data with a ratio of 80:10:10. The model training process was carried out using fine-tuning

techniques using the IndoBERT-base-p1 model. Model performance was evaluated on the test data using a confusion matrix and classification report to measure sentiment classification performance.
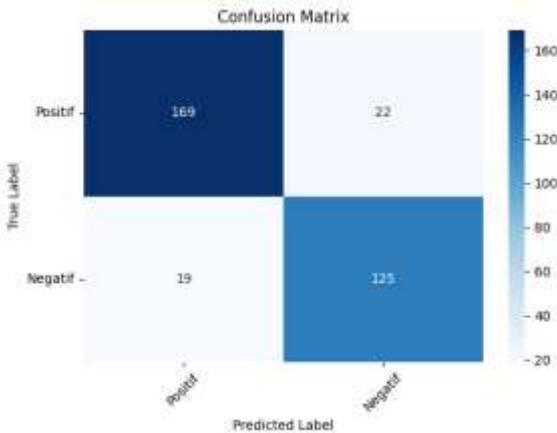

Figure 3. Confusion Matrix on Test Data

Based on the evaluation results, the IndoBERT-base-p1 model achieved an accuracy rate of 87.76%. The positive sentiment class achieved a precision of 90%, a recall of 88%, and an F1-score of 89%, while the negative sentiment class achieved a precision of 85%, a recall of 87%, and an F1-score of 86%. The high weighted average value indicates the model's good generalization ability in classifying user sentiment.


```
Accuracy: 0.8776119402985074

Classification Report:
                precision    recall  f1-score   support

      Positif        0.90      0.88      0.89       191
      Negatif        0.85      0.87      0.86       144

     accuracy                            0.88       335
    macro avg        0.87      0.88      0.88       335
 weighted avg        0.88      0.88      0.88       335
```

Figure 4. Classification Report Results

Further analysis was conducted using word cloud visualization to identify the dominant words in each sentiment class. In the positive sentiment category, words such as "good," "helpful," "route," and "schedule" appeared with high frequency. This indicates that users found the TJ: Transjakarta app to be helpful in accessing route and schedule information.

Figure 5. Positive Sentiment Word Cloud

Conversely, in negative sentiment, words like "no," "error," "slow," and "update" dominated user reviews. These words reflect technical issues users experienced, particularly after app updates, as well as issues related to transactions and account access. This finding indicates a gap between user expectations and app performance under certain conditions.

Figure 6. Negative Sentiment Word Cloud

The distribution of sentiment classes displayed in the bar chart shows that although positive sentiment is more dominant, negative sentiment still has a significant proportion and requires special attention from application managers as a basis for service evaluation.
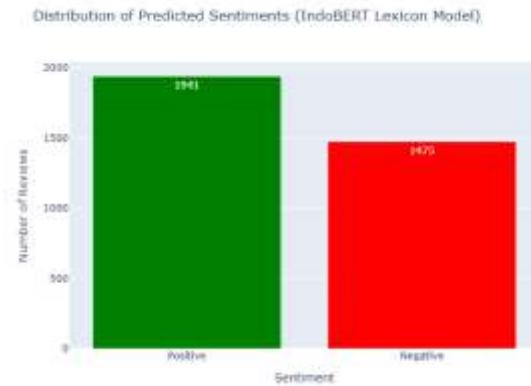
Figure 7. Bar Chart for Sentiment

Classification Results Overall, the research results show that the IndoBERT-base-p1 model is capable of classifying TJ: Transjakarta app user sentiment well. The resulting sentiment analysis provides useful insights for app managers in understanding user perceptions and determining service improvement priorities based on user complaints and needs.

CONCLUSION

This study successfully applied sentiment analysis to user reviews of the TJ: Transjakarta application using the IndoBERT-base-p1 model. Based on testing of 3,431 reviews obtained from the Google Play Store, the model was able to classify user sentiment with an accuracy rate of 87.76% and showed a balance of precision and recall values in both sentiment classes. The analysis results showed that positive user sentiment was dominated by easy access to route and schedule information, while negative sentiment was generally related to technical issues with the application, especially after system updates and transaction problems. These findings indicate that the IndoBERT model is effective for Indonesian language sentiment analysis and the research results can be used as evaluation material to improve the quality of service and user experience of the TJ: Transjakarta application.

REFERENCES
[1] F. Y. A'la, "Optimasi klasifikasi sentimen ulasan game berbahasa Indonesia: IndoBERT dan SMOTE untuk menangani ketidakseimbangan kelas," Edumatic: Jurnal Pendidikan Informatika, vol. 9, no. 1, pp. 256–265, 2025, doi: 10.29408/edumatic.v9i1.29666.
[2] A. Z. Amrullah, A. S. Anas, and M. A. J. Hidayat, "Analisis sentimen movie review menggunakan Naïve Bayes Classifier dengan seleksi fitur Chi Square," Jurnal BITe, vol. 2, no. 1, pp. 40–44, 2020, doi: 10.30812/bite.v2i1.804.

[3] J. Asher and E. P. Rachmawati, "Analisis sentimen ulasan bintang lima aplikasi Instagram di Google Play Store menggunakan IndoBERT," DINAMIK, vol. 30, pp. 318–330, 2025.

[4] G. T. Fadilah, L. Muflikhah, and R. S. Perdana, "Analisis sentimen produk hijab pada e-commerce Tokopedia menggunakan algoritma Support Vector Machine dan IndoBERT embedding," Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 9, no. 2, 2025.

[5] M. Fernanda and N. Fathoni, "Perbandingan performa labeling lexicon InSet dan VADER pada analisis sentimen Rohingya di aplikasi X dengan SVM," Jurnal Informatika dan Sains Teknologi, vol. 1, no. 3, pp. 62–76, 2024, doi: 10.62951/modem.v1i3.112.

[6] N. M. S. Hadna, P. I. Santosa, and W. W. Winarno, "Studi literatur tentang perbandingan metode untuk proses analisis sentimen di Twitter," in Proc. Seminar Nasional Teknologi Informasi dan Komunikasi (SENTIKA), Yogyakarta, Indonesia, 2016.

[7] H. Ibna Pratama and P. T. Prasetyaningrum, "Penerapan Support Vector Machine untuk analisis sentimen pada Google review hotel," Journal of Information System Research, vol. 6, no. 2, pp. 1244–1252, 2025, doi: 10.47065/josh.v6i2.6645.

[8] Kadarsih and D. Pujianto, "Application of transformer model and word embedding in sentiment analysis of Indonesian e-commerce application review," Journal of Computer Networks, Architecture and High Performance Computing, vol. 7, no. 3, pp. 720–732, 2025, doi: 10.47709/cnahpc.v7i3.6354.

[9] M. R. Manoppo et al., "Analisis sentimen publik di media sosial terhadap kenaikan PPN 12% di Indonesia menggunakan IndoBERT," Jurnal Kecerdasan Buatan dan Teknologi Informasi, vol. 4, no. 2, pp. 152–163, 2025, doi: 10.69916/jkbti.v4i2.322.

[10] W. Nurfitri and A. Chowanda, "Analisis sentimen pada kasus positif Covid-19 berdasarkan pemberitaan media di Indonesia menggunakan IndoBERT," Jurnal Ilmiah Komputer, pp. 580–593, 2024.

[11] A. S. Rizkia, W. Wufron, and F. F. Roji, "Analisis sentimen Coretax: Perbandingan pelabelan data manual, transformers-based, dan lexicon-based pada performa IndoBERT," MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. 5, no. 3, 2025, doi: 10.57152/malcom.v5i3.2151.

[12] Savitri, F. Rizki, and A. Sobri, "Implementation of BERT in sentiment analysis of national digital Samsat (SIGNAL) user reviews based on machine learning," Jurnal Ilmu Komputer dan Teknologi Informasi, pp. 67–75, 2025, doi: 10.18860/mat.v17i2.32059.

[13] U. Surapati et al., "Implementasi metode Naïve Bayes untuk mendeteksi hate speech pada Twitter," INTECOMS: Journal of Information Technology and Computer Science, vol. 6, no. 2, 2023.

[14] Y. Wu, Z. Jin, C. Shi, P. Liang, and T. Zhan, "Research on the application of deep learning-based BERT model in sentiment analysis," 2024.