Grenze International Journal of Engineering and Technology, June Issue

Sentiment Analysis for Public Transport using DeBERTa

Keya Dey[1], Rakshita Borkar[2], Ruchi Bhanushali[3] and Nataasha Raul[4]
[1-4]Sardar Patel Institute of Technology, Mumbai, India
Email: keya.dey22@spit.ac.in, rakshita.borkar22@spit.ac.in, ruchi.bhanushali22@spit.ac.in, nataasharaul@spit.ac.in

***Abstract*— The exponential growth of metro systems as a backbone for urban mobility has generated a large volume of unstructured passenger feedback online. Traditional sentiment analysis methods, which assign a single polarity to an entire review, are insufficient for this domain, as passengers often express conflicting sentiments about different aspects of the service within a single comment. This paper proposes a comprehensive framework for Aspect-Based Sentiment Analysis (ABSA) tailored for metro service reviews. A novel domain-specific dataset of 5,000+ annotated reviews is curated by collecting data from public forums across multiple Indian cities. After rigorous comparative evaluation of transformer architectures including BERT, RoBERTa, and ALBERT, the fine-tuned DeBERTav3-base-absa-v1.1 achieved strong performance with a weighted F1-score of 0.91 on the curated and labelled dataset. The disentangled attention mechanism of the model proved to be highly effective in mapping nuanced sentiments to specific aspects such as cleanliness, crowd management, and safety. The research is materialized through a full-stack web application that provides metro authorities with granular, actionable insights through sentiment analysis and enables anonymous feedback contributions from passengers. By applying ABSA, we can convert raw passenger opinions into a structured knowledge base, providing a powerful tool for data-driven public transport management.**

***Index Terms*— Aspect-Based Sentiment Analysis, ABSA, DeBERTa, Natural Language Processing, Transformer Models, Urban Mobility, Passenger Feedback, Intelligent Transportation Systems.**

I. INTRODUCTION

The rapid growth of urban populations has made metro rail systems essential for sustainable public transportation in cities such as Mumbai, Delhi, Bangalore, and Hyderabad [1] [2]. As reliance on metro services increases, passengers expect high standards of punctuality, safety, cleanliness, and comfort [3]. At the same time, commuters actively share their experiences on platforms like Google Maps, Reddit, and travel forums, creating a large volume of valuable feedback data [4] [5]. However, manually analysing this feedback is impractical due to its scale [6].

Traditional sentiment analysis methods classify entire reviews as positive, negative, or neutral, which often fails to capture the complexity of passenger opinions [7]. A single review may praise cleanliness while criticizing overcrowding, making overall sentiment labels insufficient for identifying specific service issues [8]. This lack of detail limits the ability of transit authorities to make targeted operational improvements [9] [10].

Grenze ID: 01.GIJET.12.2.563

Aspect-Based Sentiment Analysis (ABSA) addresses this limitation by identifying specific service aspects and determining the sentiment expressed toward each aspect [11] [12]. By extracting aspect-sentiment pairs, ABSA provides detailed insights into passenger experiences, enabling authorities to identify and prioritize critical service concerns more effectively [13]-[17].
The main contributions of this work are:

- Development of a novel annotated dataset containing over 5,000 metro service reviews for transportation-focused NLP research [18].
- Benchmarking of transformer-based models, including BERT, RoBERTa, ALBERT, and DeBERTa, for ABSA in the metro transport domain [19].
- Design and deployment of an end-to-end web application for passenger feedback analysis and administrative decision support [20].

The remainder of the paper is organized as follows: Section II reviews related work on sentiment analysis and ABSA. Section III explains the dataset creation and methodology. Section IV describes the system architecture and implementation. Section V presents experimental results and analysis, while Section VI concludes the paper and discusses future work.

II. Related Work

This work lies at the intersection of sentiment analysis, NLP, and transportation analytics.

A. Evolution of Sentiment Analysis

Sentiment analysis has evolved from lexicon-based methods such as SentiWordNet and VADER [5] to machine learning approaches using SVM and Naive Bayes with TF-IDF features [4]. While these methods improved classification accuracy, they struggled with contextual understanding and semantic relationships.
Deep learning models, including RNNs and LSTMs [20], enhanced sequence modelling capabilities. The introduction of Transformer-based architectures [8], especially BERT, significantly improved contextual language understanding through pre-trained embeddings. Later models such as RoBERTa, ALBERT, and DeBERTa further improved efficiency and performance on NLP tasks.

B. Aspect-Based Sentiment Analysis (ABSA)

ABSA extends traditional sentiment analysis by identifying specific aspects within text and determining sentiment toward each aspect [12]. Earlier ABSA methods used separate pipelines for aspect extraction and sentiment classification, while recent transformer-based approaches perform both tasks more effectively in an end-to-end manner. Benchmark datasets from the SemEval workshops have played a key role in advancing ABSA research.

C. Sentiment Analysis in Transportation

Sentiment analysis has been applied to domains such as airlines, hotels, and e-commerce reviews [2] [15]. However, research focused specifically on metro rail systems remains limited. Existing studies often rely on basic sentiment classification and do not provide comprehensive ABSA frameworks. This work addresses that gap through a large-scale metro review dataset, transformer-based benchmarking, and a deployable web application for metro feedback analysis.

III. Methodology

The methodology for this project was structured into four major phases: dataset creation and annotation, preprocessing and augmentation, model training and evaluation, and sentiment scoring for comparative analysis.

A. Dataset Creation and Annotation

Since no large-scale annotated dataset existed for Indian metro services, a custom corpus was created from publicly available online discussions containing passenger feedback. Reviews were collected from city-specific communities related to the metro systems of Mumbai, Delhi, Bangalore, and Hyderabad. Only transportation-related reviews were retained to ensure relevance and diversity in urban transit experiences. Nine important service aspects were identified from recurring passenger concerns: Metro Station Infrastructure, Metro Station Connectivity, General Safety, Crowd Management, Metro Frequency, Cleanliness, Women’s Safety, Ticketing System and Staff Behaviour.

An LLM-assisted annotation pipeline was used to extract aspect–sentiment pairs from reviews. Each sentiment was classified as Positive, Negative, or Neutral, followed by manual verification for consistency and accuracy. The final dataset was organized as sentence, aspect, sentiment triplets.

B. Data Preprocessing and Augmentation

The dataset underwent preprocessing steps such as lowercasing, punctuation removal, whitespace normalization, and lexical consolidation to improve consistency.

Initial analysis revealed severe class imbalance, with the neutral class containing significantly fewer samples than positive and negative classes.

To address this issue, synonym replacement augmentation was applied specifically to neutral reviews. Words were replaced with contextual synonyms while preserving sentence meaning, generating approximately 1,000 additional neutral samples. This improved class balance and helped the model better recognize nuanced or mixed feedback.

C. Model Selection and Training

The task was framed as an aspect-based sentiment classification problem, where the model predicts sentiment polarity for a given sentence and aspect. Four transformer-based models were benchmarked:

- BERT-base-uncased
- RoBERTa-base
- ALBERT-base-v2
- DeBERTa-v3-base-absa-v1.1

The dataset was split into 80% training and 20% testing using stratified sampling. Models were trained using GPU acceleration, mixed-precision training, low learning rates, and linear warmup scheduling. Performance was evaluated using Accuracy, Precision, Recall, and F1-score, with greater emphasis on weighted and macro F1-scores due to the multi-class nature of the problem.

Table 1 shows that DeBERTa achieved the best overall performance with a weighted F1-score of 0.91. Its disentangled attention mechanism and ABSA-specific pretraining allowed it to better capture relationships between aspect terms and contextual sentiment expressions. The fine-tuned DeBERTa model performs well, particularly for the neutral class, where the high recall demonstrates the effectiveness of the data augmentation strategy in improving balanced sentiment recognition.

TABLE I. COMPARATIVE PERFORMANCE OF TRANSFORMER MODELS ON METRO REVIEWS

MODEL	ACCURACY	PRECISION (WEIGHTED)	RECALL (WEIGHTED)	F1-SCORE (WEIGHTED)
BERT	0.88	0.88	0.88	0.88
ALBERT	0.82	0.82	0.82	0.82
ROBERTA	0.87	0.87	0.87	0.87
DEBERTA	0.91	0.91	0.91	0.91

D. Sentiment Scoring Module

To enable comparative analysis across cities and service aspects, textual sentiments were converted into numerical scores using the following mapping:

Positive = 100
Neutral = 50
Negative = 0

The overall sentiment score for a particular aspect or city was calculated using the average sentiment score formula:

$$\text{Sentiment Score} = \frac{(100 \times P) + (50 \times N_u) + \left(0 \times N_g\right)}{P + N_u + N_g}$$

Where:
P = Number of positive reviews
Nu = Number of neutral reviews
Ng = Number of negative reviews

This formula produces a standardized score between 0 and 100, where higher values indicate more positive public perception of metro services.

IV. SYSTEM ARCHITECTURE AND IMPLEMENTATION

The proposed research was implemented as a web-based application designed to make aspect-based sentiment analysis (ABSA) insights easily accessible and interpretable. The overall architecture is divided into two major components: the Model Development Phase and the Inference Phase, which together enable automated analysis of metro-related passenger reviews using a fine-tuned DeBERTa model.

A. System Architecture

The development phase focuses on preparing and training the sentiment analysis model. Metro-related reviews were collected from multiple online sources such as forums, websites, and mobile applications. The collected data was then cleaned and normalized through preprocessing techniques including lowercasing, removal of punctuation, URLs, emojis, and other noisy elements. After preprocessing, each review was annotated with relevant metro service aspects such as cleanliness, punctuality, safety, and fare, along with their corresponding sentiment polarity labels (Positive, Negative, or Neutral). The annotated dataset was used to fine-tune the pretrained DeBERTa model for aspect–sentiment classification. Once training was completed, the optimized model weights were saved and integrated into an inference pipeline capable of performing real-time sentiment prediction.

The inference phase handles real-time interaction with users. A user submits a metro-related review through the web interface or API, after which the system applies the same preprocessing pipeline used during training. Relevant aspects are extracted from the review, and the fine-tuned DeBERTa model predicts the sentiment associated with each aspect.

The generated aspect-wise sentiment results are then displayed to the user in an interpretable format, enabling quick understanding of passenger opinions and service quality.

B. Key Features

The user interface was designed to support both passengers and metro administrators by providing clear and actionable insights.

- Station Overview Dashboard: Provides a summarized view of selected metro stations, including overall ratings, review counts, and aspect-wise sentiment distribution using a color-coded visualization system.
- Granular Review Analysis: Displays individual reviews annotated with extracted aspects and corresponding sentiment labels, allowing detailed analysis of passenger feedback.
- Anonymous Review Submission: Enables users to submit reviews without account creation, encouraging wider participation and continuous collection of authentic passenger feedback.

V. RESULTS AND DISCUSSION

This section presents the experimental findings and demonstrates the practical application of the fine-tuned DeBERTa model on real-world metro station reviews collected from Google Maps and Reddit. The analysis focused on stations across Mumbai's Blue Line as well as metro systems in major Indian metropolitan cities.

A. Metro Station Sentiment Analysis

The deployed ABSA system successfully identified aspect-wise sentiment trends and highlighted station-specific strengths and weaknesses. Across most stations, cleanliness and infrastructure received strong positive sentiment, while crowd management consistently emerged as the most negatively perceived aspect.

B. Andheri Station Analysis

A total of 778 passenger reviews were analysed for Andheri Metro Station. Passengers showed strong approval of cleanliness and metro station infrastructure, whereas crowd management received predominantly negative feedback, indicating severe congestion issues. Concerns regarding women's safety were also observed.

C. Azad Nagar Station Analysis

Azad Nagar Station generated 372 reviews reflecting generally positive commuter experiences. Passengers highly appreciated cleanliness, infrastructure, connectivity, and metro frequency, with very limited negative sentiment across aspects.

D. Ghatkopar Station Analysis

Ghatkopar Station received the highest passenger engagement with 936 reviews due to its role as a major transit hub. While passengers praised connectivity and cleanliness, crowd management received extremely high negative sentiment because of heavy congestion during peak hours. Metro frequency also attracted criticism due to overcrowding and operational delays.

E. Saki Naka Station Analysis

Saki Naka Station also demonstrated high commuter engagement with 814 reviews. Passengers expressed strong satisfaction with station infrastructure and cleanliness, while crowd management again emerged as the most significant issue. Metro frequency and connectivity were generally viewed positively.

F. WEH Station Analysis

WEH Station reviews reflected active commuter participation and mixed operational feedback. The station performed strongly in cleanliness, infrastructure, and connectivity, receiving overwhelmingly positive responses. However, women's safety and metro frequency showed comparatively lower sentiment scores.

G. City-wise Analysis

Fig. 1 compares sentiment scores across metro systems in Mumbai, Delhi, Bangalore, and Hyderabad. Mumbai performed strongly in cleanliness but showed severe concerns regarding crowd management and general safety. Bangalore recorded poor performance in staff behavior and safety-related aspects. Delhi achieved the highest connectivity score but performed poorly in cleanliness and women's safety. Hyderabad emerged as the weakest overall, particularly in crowd management and women's safety.

Overall, the sentiment analysis reveals that while metro systems generally perform well in infrastructure and connectivity, significant concerns remain regarding crowd management, safety, and passenger comfort. These findings validate the effectiveness of the proposed ABSA system in generating targeted and actionable insights beyond traditional overall satisfaction metrics.

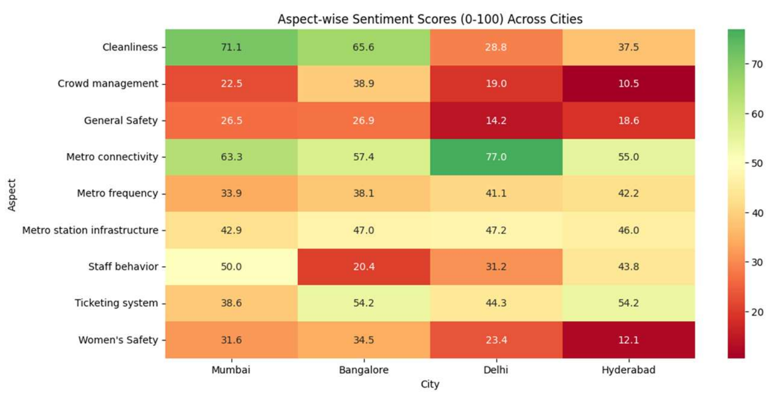

Fig. 1 Sentiment Score Distribution Across Major Metropolitan Cities

VI. Conclusion

This paper presented a complete framework for applying Aspect-Based Sentiment Analysis to the domain of metro services. We demonstrated that the DeBERTa model, finetuned on a carefully curated and augmented domain-specific dataset, achieves superior performance in classifying fine-grained passenger sentiments. The deployed system successfully transforms unstructured public feedback into structured, actionable knowledge, offering a significant upgrade over traditional survey methods or overall sentiment analysis.

Future work can focus on real-time sentiment monitoring through social media and metro apps, tracking how passenger sentiments change over time in response to events or policy changes, and performing causal analysis to identify the underlying reasons behind negative feedback. These improvements can support more responsive, data-driven, and passenger-centric metro transportation systems.

References

[1] H. Liu, I. Chatterjee, M. Zhou, X. S. Lu, and A. Abusorrah, "Aspect-Based Sentiment Analysis: A Survey of Deep Learning Methods," IEEE Transactions on Computational Social Systems, vol. 7, no. 6, pp. 1358–1375, 2020.

[2] S. Gunasekar and S. Sudhakar, "Does hotel attributes impact customer satisfaction: A sentiment analysis of online reviews," Journal of Global Scholars of Marketing Science, vol. 29, no. 2, pp. 180–195, 2019.

[3] M. Pan, N. Li, and X. Huang, “Asymmetrical impact of service attribute performance on consumer satisfaction: an asymmetric impact-attention-performance analysis,” Information Technology & Tourism, vol. 24, no. 2, pp. 221–243, 2022.
[4] G. Chauhan, A. Sharma, and N. Dwivedi, “Amazon Product Reviews Sentimental Analysis using Machine Learning,” in 2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT), 2024, pp. 1202–1206.
[5] J. C. Gope, T. Tabassum, M. M. Mabrur, K. Yu, and M. Arifuzzaman, “Sentiment Analysis of Amazon Product Reviews Using Machine Learning and Deep Learning Models,” in 2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE), 2022, pp. 1–6.
[6] R. Annamalai, S. A. Rasool, S. Deena, K. Venkatraman, and Y. Soundaram, “Sentiment Analysis using VADER: Unveiling Customer Sentiment and Predicting Buying Behavior in the Airline Industry,” in 2024 IEEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE), 2024, pp. 277–282
[7] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[8] S. K. Sahoo and S. S. Goswami, “A Comprehensive Review of Multiple Criteria Decision-Making (MCDM) Methods: Advancements, Applications, and Future Directions,” Decision Making Advances, vol. 1, no. 1, pp. 25–48, 2023.
[9] L. Zhang, S. Wang, and B. Liu, “Deep Learning for Sentiment Analysis: A Survey,” Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 8, no. 4, p. e1253, 2018.
[10] H. Liao, Y. He, X. Wu, Z. Wu, and R. Bausys, “Reimagining Multi-Criterion Decision Making by Data-Driven Methods Based on Machine Learning: A Literature Review,” Information Fusion, vol. 100, Art. no. 101970, 2023.
[11] Y. C. Hua, P. Denny, J. Wicker, and K. Taskova, “A Systematic Review of Aspect-Based Sentiment Analysis: Domains, Methods, and Trends,” Artificial Intelligence Review, vol. 57, article 296, 2024.
[12] T. Sangeetha, N. Balaganesh, and K. Muneeswaran, “Aspects based opinion mining from online reviews for product recommendation,” in Proc. 2017 Int. Conf. on Computational Intelligence in Data Science (ICCIDS), 2017, pp. 1–5.
[13] D. Jayakody, A. V. A. Malkith, K. Isuranda, V. Thenuwara, N. de Silva, S. R. Ponnamperuma, G. G. N. Sandamali, and K. L. K. Sudheera, “Instruct-DeBERTa: A Hybrid Approach for Aspect-based Sentiment Analysis on Textual Reviews,” International Journal on Advances in ICT for Emerging Regions, vol. 18, no. 2, pp. 39–50, 2025.
[14] Y. Ma, H. Peng, E. Cambria, and A. Hussain, “Targeted Aspect-Based Sentiment Analysis via Embedding Commonsense Knowledge into an Attentive LSTM,” Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32, no. 1, pp. 5876–5883, 2018.
[15] E. S. Alamoudi and N. S. Alghamdi, “Sentiment classification and aspect-based sentiment analysis on Yelp reviews using deep learning and word embeddings,” Journal of Decision Systems, vol. 30, no. 2–3, pp. 259–281, 2021.
[16] P. He, X. Liu, J. Gao, and W. Chen, “DeBERTa: Decoding-enhanced BERT with Disentangled Attention,” in International Conference on Learning Representations (ICLR), 2021.
[17] E.-S. Apostol, A.-G. Pisică, and C.-O. Truică, “ATESA-BÆRT: A Heterogeneous Ensemble Learning Model for Aspect-Based Sentiment Analysis,” Knowledge-Based Systems, vol. 326, article 113987, 2025.
[18] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, “SemEval-2014 Task 4: Aspect Based Sentiment Analysis,” in Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), 2014, pp. 27–35.
[19] K. Aziz, D. Ji, P. Chakrabarti, T. Chakrabarti, M. S. Iqbal, and R. Abbasi, “Unifying aspect-based sentiment analysis BERT and multi-layered graph convolutional networks for comprehensive sentiment dissection,” Scientific Reports, vol. 14, article 14646, 2024.
[20] B. Rahman and Maryani, “Optimizing Customer Satisfaction through Sentiment Analysis: A BERT-based Machine Learning Approach to Extract Insights,” IEEE Access, vol. 12, pp. 151476–151489, 2024.