JUMANTARA
Jurnal Manajemen dan Teknologi Rekayasa
ISSN: 2964-0741 (Print)
ISSN: 2962-7079 (Online)
DOI: 10.28989/jumantara.v4i2.3043

Listening to Student Voices: Aspect-Based Sentiment Analysis of Academic Services Using BERT

Purwo Setiawan[1,*], Arga Seta Asmara Sakti[2], Dinda Safitri Ramadhani[3]
[1,2,3] Department of Industrial Engineering, Universitas Muhammadiyah Surakarta, Indonesia

Article Info

Article history:

Received June 3, 2025
Accepted June 10, 2025
Published July 1, 2025

Keywords:

Sentiment Analysis
ABSA
BERT
Academic Services
Student Feedback

ABSTRACT

The inability to systematically process large volumes of unstructured student feedback hinders the enhancement of academic service quality in higher education. To address this challenge, this study develops an Aspect-Based Sentiment Analysis (ABSA) model using a fine-tuned BERT architecture. applied to 1,110 student reviews at Universitas Muhammadiyah Surakarta. The model was trained and evaluated using a dataset of 1,110 student reviews, filtered from an initial dataset of over 40,000 raw data points. To assess its performance, standard metrics such as accuracy, precision, recall, and F1-score were employed. The model demonstrated high performance, achieving an overall accuracy of 98.6% and an F1-score of 0.92 for identifying service aspect terms. The analysis successfully extracted key aspects, including staff interaction, administrative processes, and service efficiency. Critically, it revealed that staff interaction was the aspect with the most significant negative sentiment, providing a clear target for institutional improvement. This research confirms that the BERT-based ABSA model is a reliable and scalable tool for transforming qualitative student feedback into actionable, data-driven insights, enabling targeted enhancements to academic service quality.

Corresponding Author:

Purwo Setiawan,
Department of Industrial Engineering,
Universitas Muhammadiyah Surakarta,
Jl. A. Yani, Pabelan, Kartasura, Sukoharjo, Jawa Tengah 57169.
Email: *ps839@ums.ac.id

1. INTRODUCTION

Universitas Muhammadiyah Surakarta (UMS) is one of the largest private universities in Indonesia, graduating approximately 7,000 students annually. In an effort to enhance the quality of academic services, the Academic Administration Bureau (BAA) plays a crucial role in managing various academic administrative needs, including student registration, class schedule management, academic document archiving, and services related to graduation and commencement ceremonies. As part of service improvement efforts, every student scheduled to graduate is required to complete a feedback form containing criticisms, suggestions, or input regarding BAA services. However, to date, this data has not been optimally processed, resulting in suboptimal evaluation of the services provided. This is evident from the ongoing negative feedback from students regarding BAA services, indicating that data-driven improvements have not been fully implemented.

In recent years, sentiment analysis has become a highly valuable tool for businesses and educational institutions in understanding customer or student perceptions of the services provided. A study conducted by Liu et al. revealed that 98% of consumers read online reviews when evaluating a service, and 76% stated that negative reviews influence their decisions [1]. Supporting this, research on the impact of online reviews shows that negative feedback significantly reduces the probability of a purchase or decision [2]. This occurs because negative reviews often contrast with high average ratings, decrease the likelihood that consumers will continue to seek information, and increase the probability that they will search for substitute products or services. The effect is particularly strong when reviews pertain to functionality or customer service, a context highly relevant to academic services. In the context of higher education, research by Dervenis et al. found that over 70% of students across various universities feel that their feedback on campus services is often ignored, leading to a decline in their overall satisfaction with the institution [3].

Research on sentiment analysis has been widely conducted across various fields, including education and healthcare, with the goal of understanding users' opinions toward a service. A review study by Pooja et al. highlighted the role of sentiment analysis in improving educational quality, particularly in evaluating student opinions about their institutions through social media data [4]. Additionally, research by Ardiansyah et al. utilized a BERT model to analyze sentiment in healthcare service reviews on Google Maps, demonstrating that deep learning-based methods can identify specific aspects most frequently commented on by users [5]. On the other hand, studies on sentiment analysis in higher education have also been conducted to assess student feedback on learning, but most of them still focus on technical approaches rather than their application in academic decision-making [6].

A systematic, automated, and data-driven analysis method is required to address this issue by processing student feedback more effectively. One effective approach in sentiment analysis is Aspect-Based Sentiment Analysis (ABSA). ABSA is a technique within Natural Language Processing (NLP) that not only determines sentiment polarity (positive, negative, or neutral) but also identifies specific aspects that are the subject of the sentiment [7]. In the context of this study, ABSA enables the model to recognize specific service aspects of the Academic Administration Bureau (BAA) such as service speed, staff responsiveness, or clarity of administrative procedures and to classify students' sentiments toward these aspects.

The approach used in this study is fine-tuning a pre-trained BERT (Bidirectional Encoder Representations from Transformers) model, which has proven to be highly effective in various NLP tasks due to its ability to understand word context bidirectionally [8]. The model will be trained using the AdamW optimizer, a gradient-based optimization method designed to improve model performance in deep learning by providing more stable weight regularization compared to the classic Adam optimizer [9].

As framed by comprehensive surveys in the field, the advent of Pre-trained Language Models (PLMs) like BERT represents a paradigm shift in ABSA, fundamentally overcoming the limitations of prior techniques. The primary strength of BERT lies in its deep, bidirectional contextual understanding, allowing it to interpret a word's meaning based on the entire sentence a significant leap from the unidirectional context of earlier LSTMs and the static nature of traditional, feature-reliant machine learning. Stemming from its pre-training on massive text corpora, BERT effectively eliminates the need for laborious feature engineering and enables a powerful transfer learning approach. Consequently, by leveraging this vast, pre-existing linguistic knowledge, BERT can be fine-tuned with smaller, task-specific datasets to achieve state-of-the-art (SOTA) performance and superior generalization across nearly all ABSA subtasks [10].

The application of fine-tuning BERT in ABSA has shown significant results across various domains. In the digital application sector, a study using the IndoBERT model to analyze reviews of the M-Paspor app found that ABSA techniques can be employed to evaluate user experience more specifically, thus aiding the development of user feedback-based application services [11]. A similar study was conducted on in Drive app reviews, where BERT successfully identified key aspects of user concern, such as travel fares and driver service quality [12]. Additionally, in public opinion analysis on social media, BERT has been applied to understand public response to specific policies, such as sentiment toward the COVID-19 booster vaccine, achieving an accuracy rate of 94.81% [13]. These findings demonstrate that BERT holds significant potential in enhancing the understanding of user feedback across various sectors, including university academic services.

The novelty of this research lies in its specific application and context. While sentiment analysis has been applied in education, this study addresses an under-researched area by focusing specifically on student feedback regarding internal university administrative services managed by the Academic Administration Bureau (BAA). This approach moves beyond general reviews of teaching or facilities to analyze feedback on core operational functions. Secondly, the study utilizes a unique, non-public dataset of official feedback collected during graduation registration, offering more direct and candid insights than publicly available data from social media or review platforms.

Through this study, it is expected that a BERT-based Aspect-Based Sentiment Analysis (ABSA) model can be developed to automatically analyze student feedback on BAA services. This research aims to address the following questions:

1. How can an Aspect-Based Sentiment Analysis (ABSA) model be built to identify service aspects and sentiments contained in student feedback toward academic services?
2. How well does a BERT model fine-tuned using the AdamW optimizer perform in classifying aspects and sentiments?
3. To what extent can the model assist BAA in evaluating academic services more quickly and accurately?

With this model, BAA will be able to promptly identify service aspects that require improvement and make data-driven decisions to enhance the quality of academic services at UMS.

2. METHODOLOGY

This study uses a quantitative approach with an experimental method to develop and evaluate a BERT-based ABSA model. The data used in this study consists of student reviews related to academic services at Universitas Muhammadiyah Surakarta.

2.1. Data Sources and Processing

2.3.1. Data Sources

Data were collected from student reviews regarding academic services, including the Academic Administration Bureau (BAA), academic information systems, and other academic administrative services. The data were obtained from the graduation website database through forms filled out by students when registering for graduation. To ensure ethical standards were met, explicit consent to use this data for research purposes was obtained from the institution. Furthermore, all personal identifying information was removed from the dataset to protect student privacy, ensuring all feedback was fully anonymized before analysis.

2.3.2. Data Preprocessing Steps

Before being used for model training, the raw text data underwent several essential preprocessing steps to clean, normalize, and structure the text, thereby improving the model's effectiveness. The steps are as follows:

1. Case folding (converting all text to lowercase).
2. Text cleaning (removal of irrelevant special characters, numbers, and punctuation).
3. Tokenization (splitting text into words or phrases).
4. Stop word removal (removal of common words that do not have significant meaning).
5. Stemming and lemmatization (converting words to their root form).

2.3.3. Dataset Splitting

The data was divided into three parts: 70% for training, 20% for testing, and 10% for validation.

2.2. Model and Training Method

This study uses a fine-tuned BERT model for Aspect-Based Sentiment Analysis (ABSA). The model is trained using the AdamW optimizer with a learning rate of 5e-5, a batch size of 8, and is run for 3 epochs using Cross-Entropy Loss as the loss function. The AdamW process can be formulated as follows:

$$\theta_t = \theta_{t-1} - \eta \cdot \left(\frac{m_t}{\sqrt{v_t} + \epsilon} + \lambda \cdot \theta_{t-1} \right) \tag{1}$$

With:

θ_t : the updated parameter (weight) at iteration t
m_t, v_t : the first and second moment estimates of the gradients, respectively.
η : learning rate
λ : weight decay parameter
ϵ : a small constant to prevent division by zero

In its implementation, the BERT model will be fine-tuned using the student review dataset that has undergone the preprocessing steps, with predefined academic service aspects. The goal of this model is to classify the sentiment contained in each aspect mentioned in the reviews, providing deeper insights into student perceptions of academic services at Universitas Muhammadiyah Surakarta.

2.3. Model Evaluation

Model evaluation is performed using several performance metrics, including:

2.3.1. Precision

Precision is useful for measuring the extent to which the aspect tokens predicted by the model are actually correct aspect tokens, thus indicating how accurate the model's predictions are. The precision formula can be written as:

$$Precision = \frac{True\ Positive}{(True\ Positives + False\ Positives)} \tag{2}$$

True Positives (TP): The number of aspect tokens correctly predicted by the model
False Positives (FP): The number of non-aspect tokens incorrectly predicted as aspect tokens

2.3.2. Recall

Recall is useful for measuring the extent to which the model is able to identify the aspect tokens present in the data. In this case, recall reflects the model's sensitivity in recognizing aspect tokens. The recall formula can be written as:

$$Recall = \frac{True\ Positive}{(True\ Positives + False\ Negatives)} \tag{3}$$

True Positives (TP): The number of aspect tokens correctly predicted.
False Negatives (FN): The number of actual aspect tokens that the model failed to predict.
2.3.3. F1-score:

F1-score is used to evaluate the overall performance of the model in recognizing aspect tokens, especially when it is necessary to balance between precision and recall. The F1-score formula can be written as:

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

2.3.4. Accuracy

Accuracy is used to measure the proportion of all tokens (both aspect and non-aspect tokens) that are correctly classified by the model. Thus, a high accuracy indicates that the model is generally capable of classifying tokens correctly. The accuracy formula can be written as:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Samples}} \tag{5}$$

3. RESULT AND DISCUSSION

In this study, the implementation was carried out using Python on the Google Colab platform. To initiate a Python program in Google Colab, it is necessary to use libraries collections of pre-written code that can be reused across multiple programs. For conducting sentiment analysis in Python, several essential libraries include:

- NumPy: Numerical computations.
- Pandas: Data manipulation and analysis.
- spaCy: Text preprocessing and tokenization.
- TextBlob: Sentiment analysis.
- PyTorch/TensorFlow: Deep learning framework for model development.
- Hugging Face Transformers: Pre-trained transformer models (using BERT).
- Matplotlib/Seaborn: Data visualization.
- SHAP/LIME: Explainability for sentiment model.

These libraries can be installed using the !pip install command, and once installed, they are ready to be used by importing them with the import statement.

3.1. Data Collection

This study uses data from 40,000 raw data points. The data were obtained from the graduation website database through forms filled out by students when registering for graduation.

3.2. Data Preprocessing

The data preprocessing stage began by filtering an initial dataset of over 40,000 raw data points to create a refined corpus of 1,110 review entries suitable for analysis. This initial filtering was crucial for data quality and involved two criteria: eliminating duplicate sentences and retaining only those reviews containing five or more words. After this selection, the reviews underwent a standard preprocessing pipeline to clean and normalize the text, which included case folding, text cleaning, tokenization, stop word removal, and stemming/lemmatization. Table 1 presents examples of the data after undergoing these processes.

Table 1. Preprocessing Result

Process	Text	
	Before	After
Case folding	True, BAA Officer annoying, Unfriendly. BAA Officer student unfriendly, make heart irritated	true, baa officer annoying, unfriendly. baa officer student unfriendly, make heart irritated
Text cleaning	true, baa officer annoying, unfriendly. baa officer student unfriendly, make heart irritated	true baa officer annoying unfriendly baa officer student unfriendly make heart irritated
Case Folding	true baa officer annoying unfriendly baa officer student unfriendly make heart irritated	['true', 'baa', 'officer', 'annoying', 'unfriendly', 'baa', 'officer', 'student', 'unfriendly', 'make', 'heart', 'irritated']

Tokenization	['true', 'baa', 'officer', 'annoying', 'unfriendly', 'baa', 'officer', 'student', 'unfriendly', 'make', 'heart', 'irritated']	['baa', 'officer', 'annoying', 'unfriendly', 'baa', 'officer', 'student', 'unfriendly', 'heart', 'irritated']
Stop-word Removal	['true', 'baa', 'officer', 'annoying', 'unfriendly', 'baa', 'officer', 'student', 'unfriendly', 'make', 'heart', 'irritated']	['baa', 'officer', 'annoying', 'unfriendly', 'baa', 'officer', 'student', 'unfriendly', 'heart', 'irritated']
Stemming / Lemmatization	['baa', 'officer', 'annoying', 'unfriendly', 'baa', 'officer', 'student', 'unfriendly', 'heart', 'irritated']	['baa', 'officer', 'annoy', 'unfriend', 'baa', 'officer', 'student', 'unfriend', 'heart', 'irritate']

3.3. Data Labeling and Splitting

The data labeling process was carried out manually by the researchers and involved three main columns: sentence, aspect, and sentiment. Each sentence could contain more than one aspect, depending on the content and context of the review. The manual labeling process was applied to all 1,110 review entries. This fully labeled dataset was then split into three parts: 70% for training, 20% for testing, and 10% for validation. This annotation process aimed to identify relevant aspects and determine the sentiment orientation (positive or negative) expressed in each sentence. Examples of the labeled data are presented in Table 2.

Table 2. Data Labeling

Text	Sentiment	Aspect & Keywords
thank ums baa staff helping serve student academic future hopefully officer even friendlier student	Positive	{'aspect': 'staff_interaction', 'sentiment': 'positive', 'keywords': 'friendlier' }
assalamualaikum apologize advance lady gentleman spot due respect regret baa example managing personal biodata	Negative	{'aspect': 'technical_systems', 'sentiment': 'negative', 'keywords': 'error'}, {'aspect': 'information_accessibility', 'keywords': 'update'}
process managing academic files administration clear enough sometimes ineffective especially practical coass student	Negative	{'aspect': 'administrative_process', 'sentiment': 'negative', 'keywords': 'process'}
sometimes enter baa confused take care information center name tag officer sometimes bit confusing officer take care	Negative	{'aspect': 'information_accessibility', 'sentiment': 'negative', 'keywords': 'information'}
bismillahirrahmanirrahim overall good need clarify regarding notification announcement conveyed thank	Positive	{{'aspect': 'information_accessibility', 'sentiment': 'positive', 'keywords': 'notification'}, {'aspect': 'information_accessibility', 'sentiment': 'positive', 'keywords': 'announcement'}
feel service baa unfriendly student make complaint question responded well often speak high tone student often find	Negative	{'aspect': 'staff_interaction', 'sentiment': 'negative', 'keywords': 'unfriendly'}, {'aspect': 'service_efficiency', 'sentiment': 'negative', 'keywords': 'time'}

3.4. Model Performance Evaluation and Discussion

After completing the training process using a fine-tuned BERT model with the AdamW optimizer, the performance of the model was evaluated using accuracy, precision, recall, and F1-score. The evaluation results show that the model achieved a high overall accuracy of 98.6%, indicating that the model is highly reliable in classifying tokens correctly across both aspect and non-aspect categories. For aspect tokens, the model achieved: Precision: 0.87, Recall: 0.97 and F1-score: 0.92.

These scores indicate that the model is highly sensitive in identifying relevant aspect terms from student feedback while maintaining good precision. It demonstrates the model's effectiveness in recognizing specific topics or concerns mentioned in student reviews. For non-aspect tokens, the model performed exceptionally well, achieving: Precision: 1.00; Recall: 0.99 and F1-score: 0.99. This implies the model is highly accurate in excluding irrelevant tokens, ensuring that only relevant information is considered during aspect extraction. The macro-average F1-score was 0.95, and the weighted average F1-score was 0.99, both of which reflect strong overall model performance.

These results suggest that the BERT-based ABSA model is capable of accurately extracting and classifying sentiment from diverse student expressions. Several common aspects identified in the student reviews include staff friendliness, clarity of information, efficiency of service, and technical system issues. For example, reviews mentioning "slow response", "confusing procedures", or "friendly staff" were accurately tagged with appropriate aspect labels and sentiment values. This result corroborates the findings of Husnah and Yuliamir, who demonstrated the critical importance of aspects such as responsiveness and friendliness in influencing user satisfaction [14].

The application of this model provides a scalable and objective tool for higher education institutions to process qualitative student feedback, which is often left unanalyzed due to time and resource constraints. Furthermore, such analysis enables the Academic Administration Bureau (BAA) to gain insights into areas requiring service improvement, aligned with student expectations and experiences.

The findings confirm that deep learning approaches like BERT can serve as effective and efficient tools to support decision-making processes in campus service improvements. Moreover, the study emphasizes the importance of listening to student voices as a key indicator of service quality. Future work should focus on expanding the dataset to include more diverse and representative samples and integrating more advanced NLP techniques, such as attention-based models or semantic-aware approaches, to better handle implicit aspects and ambiguous expressions.

When compared to previous studies utilizing ABSA in domains such as app reviews [15] and healthcare services [16], the model in this study demonstrated competitive performance. A key strength of this research lies in its application within the higher education context in Indonesia, an area that remains relatively underexplored. However, some limitations were observed, particularly the model's difficulty in identifying implicit aspects and in interpreting ambiguous expressions. This indicates that while the model performs well quantitatively, there is room for improvement in understanding more nuanced contextual language. The model's limitation in identifying implicit aspects, as observed in this study, is a widely recognized challenge in the field of ABSA, which opens avenues for future research integrating more context-aware modules [17], [18].

3.5. Sentiment Distribution by Aspect

To further understand how students perceive each identified aspect, sentiment distribution was analyzed and visualized in Figure Y. The chart illustrates the count of positive and negative sentiments for each aspect category. Figure 1. presents the Distribution of Positive and Negative Sentiment by Aspect.

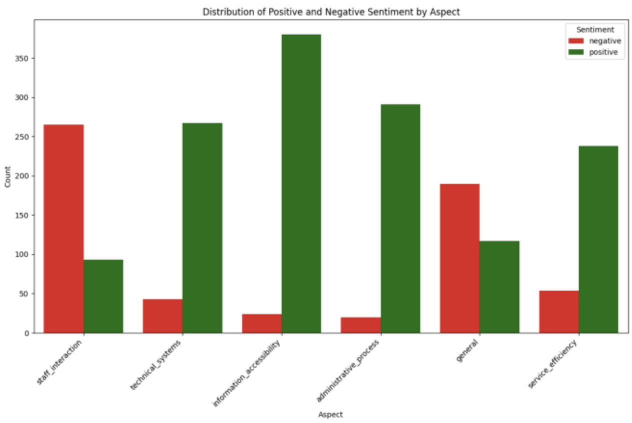

Figure 1. Distribution of Positive and Negative Sentiment by Aspect

The most notable finding is that staff_interaction received the highest number of negative sentiments, significantly outweighing the positive ones. This suggests that many students were dissatisfied with how they were treated by staff, indicating a potential area for service quality improvement, particularly in terms of interpersonal communication and professionalism. On the other hand, information_accessibility, administrative_process, and service_efficiency were predominantly associated with positive sentiments. This suggests that most students felt informed and were satisfied with the clarity and efficiency of the services provided. Technical systems also received mostly positive feedback, although it had a smaller total volume compared to other aspects.

Interestingly, the general category shows a more balanced distribution, indicating mixed feelings where some reviews express praise while others express dissatisfaction in a broad or non-specific manner. This sentiment analysis by aspect allows stakeholders to prioritize interventions not only based on frequency but also on emotional polarity. It highlights staff interaction as the most critical issue needing immediate attention, while affirming strengths in information dissemination and technical service delivery.

3.6. Practical Implications for Institutional Improvement

The most immediate and impactful implications are for the Academic Administration Bureau (BAA) at UMS. The model provides a diagnostic tool that transforms thousands of qualitative reviews from unstructured text into a clear, prioritized agenda for action.

1. Targeted Intervention for Staff Development
   The standout finding that staff_interaction receives the highest volume of negative sentiment provides an unambiguous mandate for intervention. Instead of general calls for "better service," the BAA can now implement specific, data-driven initiatives. These could include targeted professional development workshops for front-line staff focusing on empathy, active listening, and conflict resolution skills. The granular nature of the data could even pinpoint specific recurring issues (e.g., perceptions of unfriendliness or dismissiveness) that can be directly addressed in training modules.
2. Reinforcement of Strengths and Strategic Resource Allocation
   Conversely, the predominantly positive sentiment associated with information_accessibility, administrative_process, and technical_systems is equally insightful. This affirms that investments and

 ISSN : 2964-0741 (Print), 2962-7079 (Online)

strategies in these areas are effective and valued by students. The implication is not to become complacent, but to identify these successful practices and potentially apply their principles to weaker areas. For instance, the clarity and efficiency of the "administrative process" could serve as a model for improving the clarity of communication during staff-student interactions.

3. A New Paradigm for Understanding Student Experience
   Traditionally, universities rely on surveys with Likert scales, which measure satisfaction but often fail to capture the "why" behind the scores. This research demonstrates the value of analyzing unstructured, qualitative data at scale. It offers a method for other institutions to gain a deeper, more nuanced understanding of the student experience, moving beyond quantitative metrics to the rich details of student narratives.

4. CONCLUSION

This study successfully developed an Aspect-Based Sentiment Analysis (ABSA) model using a fine-tuned BERT architecture to analyze student feedback on academic services at Universitas Muhammadiyah Surakarta (UMS). By leveraging a dataset of 1,110 preprocessed and annotated student reviews, the model demonstrated strong performance in identifying both aspect terms and their associated sentiments, achieving an overall accuracy of 98.6%. The evaluation results confirmed that the model is highly capable of distinguishing relevant service aspects such as staff interaction, administrative processes, and information clarity from general text. This capability enables educational institutions to transform large volumes of qualitative feedback into actionable insights in a systematic, objective, and scalable manner.

The proposed approach offers significant potential in enhancing academic service evaluation processes by automating sentiment extraction and facilitating evidence-based decision-making. Moreover, the flexibility of the model makes it adaptable to other institutional contexts, further broadening its impact within the domain of educational quality assurance.

Despite its promising performance, the study also acknowledges certain limitations, including the model's difficulty in handling implicit sentiments and ambiguous expressions. Future work should explore the integration of larger and more diverse datasets, improved aspect detection mechanisms, and context-aware sentiment classification to further refine the model's effectiveness.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Department of Industrial Engineering, Universitas Muhammadiyah Surakarta, for the support and opportunity to conduct this research. Special thanks are extended to the Academic Administration Bureau (BAA) for granting access to student feedback data used in this study.

DAFTAR PUSTAKA (10 PT)

[1] B. Liu, *Sentiment Analysis and Opinion Mining*. Springer International Publishing, 2012. doi: 10.1007/978-3-031-02145-9.

[2] M. Varga and P. Albuquerque, "The Impact of Negative Reviews on Online Search and Purchase Decisions," *Journal of Marketing Research*, Oct. 2023, doi: 10.1177/00222437231190874/SUPPL_FILE/SJ-PDF-1-MRJ-10.1177_00222437231190874.PDF.

[3] C. Dervenis, G. Kanakis, and P. Fitsilis, "Sentiment analysis of student feedback: A comparative study employing lexicon and machine learning techniques," *Studies in Educational Evaluation*, vol. 83, Dec. 2024, doi: 10.1016/j.stueduc.2024.101406.

[4] Pooja and R. Bhalla, "A Review Paper on the Role of Sentiment Analysis in Quality Education," *SN Comput Sci*, vol. 3, no. 6, pp. 1–9, Nov. 2022, doi: 10.1007/S42979-022-01366-9/FIGURES/3.

[5] Y. Liu *et al.*, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," Jul. 2019.

[6] T. Shaik, X. Tao, C. Dann, H. Xie, Y. Li, and L. Galligan, "Sentiment analysis and opinion mining on educational data: A survey," *Natural Language Processing Journal*, vol. 2, p. 100003, Mar. 2023, doi: 10.1016/J.NLP.2022.100003.

[7] T. Shaik *et al.*, "A Review of the Trends and Challenges in Adopting Natural Language Processing Methods for Education Feedback Analysis," *IEEE Access*, vol. 10, pp. 56720–56739, 2022, doi: 10.1109/ACCESS.2022.3177752.

[8] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, Association for Computational Linguistics (ACL), 2019, pp. 4171–4186. doi: https://doi.org/10.48550/arXiv.1810.04805.

[9] I. Loshchilov and F. Hutter, “Decoupled Weight Decay Regularization,” *7th International Conference on Learning Representations, ICLR 2019*, Nov. 2017, Accessed: Jan. 31, 2025. [Online]. Available: https://arxiv.org/abs/1711.05101v3

[10] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam, “A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges,” *IEEE Trans Knowl Data Eng*, vol. 35, no. 11, pp. 11019–11038, Nov. 2023, doi: 10.1109/TKDE.2022.3230975.

[11] M. Widiansyah, F. Frazna Az-Zahra, and A. Pambudi, “Fine-Tuning Model Indobert (Indonesian Bidirectional Encoder Representations from Transformers) untuk Analisis Sentimen Berbasis Aspek pada Aplikasi M-Paspor,” *Joutica*, vol. 9, no. 2, pp. 183–195, Sep. 2024, doi: 10.30736/INFORMATIKA.V9I2.1310.

[12] D. K. Alfiki Astutik, A. Indrasetianingsih, and F. Fitriani, “Penerapan Text Mining pada Analisis Sentimen Pengguna Twitter Layanan Transportasi Online Menggunakan Metode Density Based Spatial Clustering of Applications With Noise (DBSCAN) dan K-Means,” *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, vol. 15, no. 1, Jul. 2022, doi: 10.36456/JSTAT.VOL15.NO1.A5983.

[13] I. W. Nurdian, D. O. Nooryawati, I. Yulfrian, B. P. Nevista, N. Yudistira, and R. S. Perdana, “COVID-19 Vaccines in Indonesia: A Public Opinion Study Based on Twitter Data Multidimensional Sentiment Analysis,” *ACM International Conference Proceeding Series*, pp. 165–175, Nov. 2022, doi: 10.1145/3568231.3568246.

[14] C. Husnah and H. Yuliamir, “Analysis of Sales Staff Work Ability and Service Quality on Guest Satisfaction at Openaire Resto Bar and Market,” *Journal of Education, Humaniora and Social Sciences (JEHSS)*, vol. 7, no. 3, pp. 1151–1159, Feb. 2025, doi: 10.34007/JEHSS.V7I3.2500.

[15] V. Ballas, K. Michalakis, G. Alexandridis, and G. Caridakis, “Automating Mobile App Review User Feedback with Aspect-Based Sentiment Analysis,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 14738 LNCS, pp. 179–193, 2024, doi: 10.1007/978-3-031-60487-4_14.

[16] L. Gui and Y. He, “Understanding patient reviews with minimum supervision,” *Artif Intell Med*, vol. 120, p. 102160, Oct. 2021, doi: 10.1016/J.ARTMED.2021.102160.

[17] A. Adak, B. Pradhan, and N. Shukla, “Sentiment Analysis of Customer Reviews of Food Delivery Services Using Deep Learning and Explainable Artificial Intelligence: Systematic Review,” *Foods 2022, Vol. 11, Page 1500*, vol. 11, no. 10, p. 1500, May 2022, doi: 10.3390/FOODS11101500.

[18] B. Liu, T. Lin, and M. Li, “Enhancing aspect-category sentiment analysis via syntactic data augmentation and knowledge enhancement,” *Knowl Based Syst*, vol. 264, p. 110339, Mar. 2023, doi: 10.1016/J.KNOSYS.2023.110339.

 ISSN : 2964-0741 (Print), 2962-7079 (Online)