INJECT (Interdisciplinary Journal of Communication)
p-ISSN: 2548-5857; e-ISSN: 2548-7124
Vol. 10. No. 2 December 2025: p. 2197-2214
DOI: https://doi.org/10.18326/inject.v10i2.6046

Understanding Multidimensional Patient Feedback as Healthcare Communication: An Interdisciplinary Computational Analysis

Aris Rifki Setiya Bahari*
Magister Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto, 53127, Indonesia
arisrifkisb11@gmail.com

Berlilana
Informatika, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto, 53127, Indonesia
berli@amikompurwokerto.ac.id

Fandy Setyo Utomo
Informatika, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto, 53127, Indonesia
fandy_setyo_utomo@ami-kompurwokerto.ac.id

Abstract
Patient feedback represents an important form of healthcare communication through which patients articulate experiences, evaluations, and expectations toward healthcare services. In practice, this communication is often conveyed through unstructured, subjective, and multidimensional narratives, in which a single message may simultaneously address multiple service aspects. Such characteristics complicate the systematic interpretation of patient communication, particularly when sentimental expressions are unevenly distributed and dominated by positive evaluations. This study aims to examine patient feedback as a communicative practice in healthcare by analyzing multidimensional sentiment expressions from an interdisciplinary communication perspective. Computational methods are not positioned as the primary contribution of this study, but are employed as analytical tools to support the interpretation of large-scale patient communication data. An aspect-based sentiment analysis framework with a multilabel classification scheme is used to capture how sentiments are communicated toward predefined service aspects. The dataset consists of 1,131 anonymized patient feedback texts collected from JIH Purwokerto Hospital. To reduce interpretive bias caused by imbalanced sentiment distributions that may obscure less explicit communication

expressions, label-based data balancing strategies are applied. Indonesian language modeling is used to accommodate the informal and context-dependent characteristics of patient narratives. The findings indicate that this approach enables a more structured reading of patient communication across service aspects, particularly in identifying explicit positive and negative sentiments. In contrast, neutral sentiment remains more difficult to identify due to its implicit and contextual nature, reflecting the complexity of patient communication strategies. Overall, this study demonstrates that computational analysis can function as a supportive instrument in healthcare communication research for systematically mapping multidimensional patient feedback, provided that the results are interpreted contextually rather than mechanically.

Keywords: *S*entiment Analysis, Multilabel Classification, Patient Feedback, Healthcare Communication, Computational Approach.

1. Introduction

Patient feedback has an important role in evaluating the quality of healthcare services. However, the unstructured nature of such feedback poses considerable challenges. Because feedback is often delivered in informal language, it contains subjective expressions and sentiments towards more than one aspect of the service in a single statement. This condition complicates the process of identifying communication patterns and has the potential to result in unbalanced service assessments if analyzed manually or with a less systematic approach.

In addition, imbalances in the distribution of sentiment, especially the lack of representation of neutral and negative opinions, often lead to biased classification results. This imbalance can result in biased classification outcomes and limit the effectiveness of automated sentiment monitoring systems in hospital settings. Previous research indicates that the complexity of language and the imbalance of the distribution of sentiment in patient feedback can obscure the understanding of healthcare shortages (Alkhnbashi et al., 2024; Imaduddin et al., 2023; Setiawan, 2024). Understanding patient feedback is therefore essential not only for measuring service quality but also for interpreting communication dynamics between patients and healthcare providers.

In the study of healthcare communication, patient feedback is understood as a form of public communication that reflects patients' experiences, satisfaction, and complaints about their interactions with healthcare institutions. However, patient feedback is generally delivered in the form of an unstructured free narrative, uses informal and contextual language, and often contains more than one aspect of the service in a single statement. This condition makes it difficult to interpret meaning and utilize feedback systematically, especially when the distribution of sentiment is unbalanced and tends to be dominated by positive expressions. A study by Wong et al. (2020) shows that while patient feedback has great potential to support healthcare quality and communication improvements, language complexity and data irregularities often hinder the identification of specific and effectively actionable service issues.

The limitations in interpreting patient feedback, as shown by Wong et al. (2020) Show that the main problem does not lie in the absence of data, but in the difficulty of reading and organizing the meaning of patient communication in a systematic manner. When feedback is presented in the form of complex and unbalanced free narratives, conventional evaluation approaches based on manual reading or simple categorization become inadequate. Therefore, an analytical approach is needed that can capture the diversity of language expressions, map sentiment towards various aspects of services, and manage large volumes of textual data consistently and measurably.

From an interdisciplinary communication perspective, the importance of this study lies in its effort to bridge methodological gaps between healthcare communication and computational text analysis. Patient feedback is not merely evaluative data but a communicative practice through which patients negotiate meaning, express expectations, and manage their relationship with healthcare institutions. When such communication is reduced to aggregate satisfaction scores or mechanically interpreted sentiment labels, significant nuances of patient meaning-making may be overlooked. Therefore, a systematic yet communication-sensitive analytical approach is needed to capture how patients articulate experiences across multiple service dimensions, particularly in large-scale institutional settings where manual interpretation is impractical. By combining computational analysis with a communication-oriented interpretation framework, this study contributes to the interdisciplinary development of healthcare communication research by offering a scalable method to interpret patient voices without detaching them from their communicative context.

With the advancement of text analysis technologies, Natural Language Processing (NLP) approaches have increasingly been applied to examine large-scale public communication data. Pre-trained transformer-based language models, such as BERT and its Indonesian adaptation IndoBERT, have demonstrated strong performance in various text analysis tasks, including consumer reviews, educational feedback, and healthcare-related applications (Aryanti et al., 2025; Ihtada et al., 2025; Jazuli et al., 2023). IndoBERT is trained from scratch using the Indo4B corpus, which consists of approximately four billion words collected from diverse public sources, enabling the model to capture the grammatical structure, semantics, and pragmatic context of the Indonesian language (Rohman & Agung, 2025; Wilie et al., 2020). These characteristics make IndoBERT well-suited for analyzing the informal and context-rich language commonly found in patient feedback.

To enable fine-grained sentiment analysis, this study adopts Aspect-Based Sentiment Analysis (ABSA), an approach that focuses on identifying sentiment polarity toward specific entities or attributes within a text, referred to as aspects. These aspects can be grouped into more general categories, such as services, facilities, and patient meals (Nazir & Rao, 2022). In practice, ABSA can be implemented using rule-based methods, supervised classification with predefined aspects, or sequence-labeling approaches based on neural models. For patient feedback analysis, supervised classification with predefined aspects is particularly suitable, as it aligns with service evaluation frameworks and policy-driven performance indicators (Phan & Ogunbona, 2020).

Research on patient feedback often does not adequately address the multidimensional nature of communication reflected in multilabel sentiment expressions, nor the problem of sentiment imbalance across service aspects. Furthermore, studies that explicitly integrate Indonesian pre-trained language models with data-balancing strategies within a multilabel ABSA framework for healthcare data remain scarce (Alkhnbashi et al., 2024; Mei et al., 2024). Although data balancing techniques such as the Synthetic Minority Oversampling Technique (SMOTE) have been widely used to mitigate class imbalance, their application and evaluation in Indonesian multilabel ABSA for healthcare contexts are still limited.

Based on these research gaps, this study positions a computational approach as an analytical tool for understanding patient communication patterns through textual feedback. The objective of this research is to analyze multidimensional sentiment in healthcare patient feedback using an Aspect-Based Sentiment Analysis framework with a multilabel classification scheme. The proposed approach is applied to patient feedback data from JIH Purwokerto Hospital, covering three key service aspects: services, facilities, and patient meals. By integrating language modeling and data balancing strategies, this study is expected to provide a more balanced and accurate picture of patient communication expressions and support the evaluation of healthcare communication based on empirical evidence.

2. Method

This study uses an aspect-based sentiment analysis approach to analyze patient feedback as a form of healthcare communication. In this study, service aspects are predefined based on their relationship to hospital operations, namely services, facilities, and patient food. This approach does not aim to automatically extract aspects from the text, but rather to classify the sentiments associated with each predefined aspect. Therefore, this study does not claim to perform joint aspect extraction and sentiment detection but focuses on the classification of sentiment per aspect in a multilabel scheme.

Each patient feedback entry can contain more than one aspect of the service, so a single text allows for multiple aspects and sentiment labels simultaneously. To support this, the annotation process is carried out by assigning a sentiment label (positive, negative, or neutral) to each aspect that appears in a single feedback. This annotation scheme is designed to represent the multidimensional nature of patient communication and reflects a service experience that is not always singular or linear.

Figure 1 illustrates the complete methodological path applied in this study, from raw data collection to final model evaluation. The process includes text pre-processing, multilabel coding, aspect-based SMOTE sampling, IndoBERT refinement, and model evaluation using standard classification metrics. This workflow is designed to address data imbalances while enabling aspect-aware sentiment prediction in a multilabel setting.

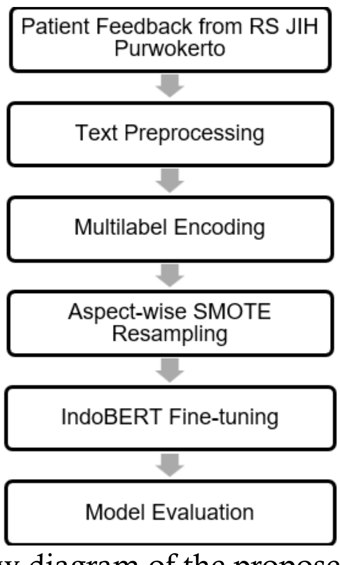

Figure 1. Workflow diagram of the proposed multilabel ABSA pipeline using IndoBERT and SMOTE

Patient feedback data show an imbalance in sentiment distribution, with the predominance of positive sentiment and limited representation of neutral and negative sentiment. To reduce the potential for bias in the model learning process, this study applied a label-wise balancing strategy. SMOTE is used separately on each sentiment label for each aspect, with the aim of improving the representation of minority classes without changing the structure of predetermined aspects.

The application of SMOTE in the context of multilabels has limitations, especially regarding the possibility of forming synthetic label combinations that do not fully reflect the natural distribution of data. Therefore, this approach is positioned as a pragmatic strategy to reduce class bias, rather than as a perfect methodological solution. These limitations are considered in the interpretation of the results and are discussed further in the discussion section.

For the analysis of Indonesian texts, this study utilizes the IndoBERT pre-trained language model as a computational modeling tool. The model is adjusted through a fine-tuning process to support multilabel sentiment classification with a multi-output approach, using a sigmoid activation function on the output layer. The training process is carried out by dividing the data into training and testing sets and using consistent training parameters to ensure the stability of the results. This computational approach is used as an analytical tool to read sentiment patterns in patient communication, rather than as the main focus of the research's methodological contribution.

2.1. Dataset

The data set used in this study consisted of 1,131 lines of anonymous patient feedback from JIH Purwokerto Hospital, which consisted of free text columns containing feedback (kritiksaran) and sentiment labels for three predetermined aspects: service (pelayanan), facility (fasilitas), and patient meals (makanpasien). Each aspect is independently labeled with one of three classes of sentiment: positive, neutral, or negative. Because a single comment can express sentiment towards various aspects, the classification task is designed as a multilabel problem.

Preliminary analysis revealed substantial class imbalances in all three aspects. For example, in the service aspect, 873 entries were labeled positive, while only 71 were neutral and 187 were negative. Similar trends are emerging in other aspects. This imbalance is shown in Table 1.

Table 1. Sentiment Distribution SMOTE

Aspect	Positive	Neutral	Negative
Pelayanan	873	71	187

Aspect	Positive	Neutral	Negative
Fasilitas	800	89	242
Makanpasien	674	98	359

2.2. Data Collection

Data was collected from the official hospital discharge form and an internal survey held at JIH Purwokerto Hospital. All patient identities are anonymized to maintain privacy. Sentiment annotation is done manually by trained healthcare staff using aspect-specific labeling guidelines. The labels are then binarized using the MultiLabelBinarizer of scikit-learn, converting the categorical sentiment into a multi-label binary vector suitable for multi-output classification tasks.

2.3. Data Processing

All textual data is pre-processed in lowercase letters, removes special characters, and applies tokenization using the IndoBERT tokenizer from the indobenchmark/indobert-base-p1 model. Sentiment labels are compiled into a multilabel binary format. To address significant imbalances in the dataset, SMOTE is applied separately to the label columns of each aspect.

After applying SMOTE per label, each sentiment class in each aspect is up-sampled to match the number of majority classes. The final distribution achieved a perfect balance with 429 samples for each class per aspect, as shown in Table 2. This balanced dataset is then recombined while maintaining the alignment of the original text, ensuring the consistency of multilabel assignments.

Table 2. Sentiment Distribution SMOTE

Aspect	Positive	Neutral	Negative	Total per Aspect
Pelayanan	429	429	429	1,287
Fasilitas	429	429	429	1,287
Makanpasien	429	429	429	1,287

2.4. Statistical Analysis

The classification model was built using the IndoBERT transformer, fine-tuned for multilabel sequence classification. The architecture utilized sigmoid activation for independent label prediction and binary cross-entropy loss. The dataset was split into training and test sets using an 80:20 ratio. The HuggingFace Trainer API was employed with three training epochs, a batch size of 8, and a maximum token length of 128. Evaluation was conducted using precision, recall, F1-score, and ROC-AUC. The final model achieved strong performance across most classes, particularly for positive and negative sentiments.

3. Results

The results showed that the computational approach used was able to classify patients' sentiments towards various aspects of healthcare services in a multilabel scheme. To assess the impact of the data balancing strategy, the performance of the model trained with label-based balancing was compared to the baseline, i.e., the model trained without the implementation of SMOTE. This comparison shows that models with data balancing result in a more representative distribution of predictions in neutral and negative sentiment classes, which were previously underrepresented.

After the implementation of the data balancing strategy and language model adjustment, the classifiers achieved a micro mean F1 score of 0.83 and a macro average F1 score of 0.83. This value suggests that the model can maintain relatively consistent performance across sentiment labels, even though preliminary data show an imbalance in class distribution. However, this result is not interpreted as achieving a completely balanced distribution condition, but rather as an indication of a reduction in classification bias against the majority class.

The performance of the models in each sentiment class is presented in Table 3. Positive sentiment shows the highest F1 score (0.93), followed by negative sentiment (0.88), while neutral sentiment achieves an F1 score of 0.69. This pattern is in line with findings in previous sentiment classification research, where neutral expressions tend to be more difficult to identify consistently due to their ambiguous and contextual nature.

3.1. Classification Report

To provide a clearer picture of the model's performance in each sentiment class, the classification results are presented in the form of a heatmap. This visualization is used to show the comparison of precision, recall, and F1 scores on each sentiment label in a multilabel classification scheme, so that performance differences between classes can be observed more intuitively before being discussed in detail.

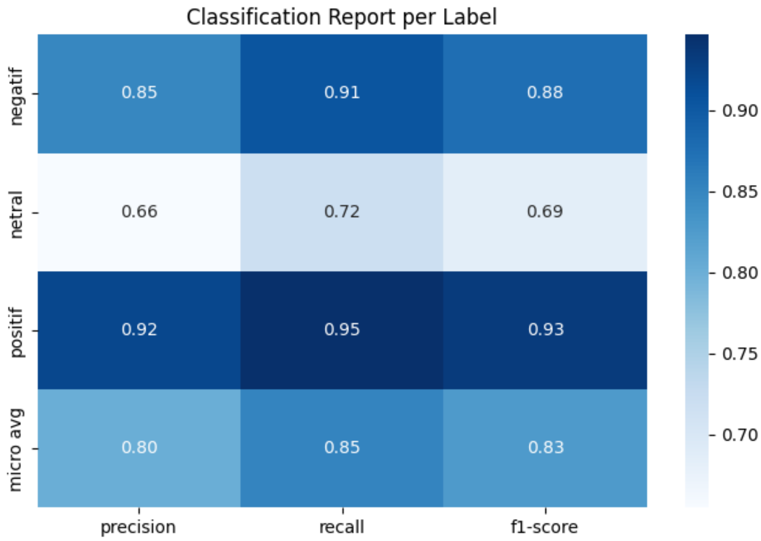

Figure 2. Summary of classification results (after SMOTE). Values reflect strong model performance in positive and negative classes.

Figure 2. display classification reports in the form of heat maps that summarize the performance of IndoBERT-based multilabel sentiment classifiers after the implementation of data balancing strategies using SMOTE. Each row represents a sentiment class, while each column shows the main evaluation metrics, namely precision, recall, and F1 score. This visualization makes it easy to identify performance differences between sentiment classes comparatively.

The results on the heatmap show that positive sentiment is the best-performing class, with a precision value of 0.92, a recall of 0.95, and an F1 score of 0.93. A high recall value indicates that most of the expressions of positive sentiment are correctly recognized by the model. These findings suggest that affirmative expressions in patient feedback tend to have more explicit and consistent linguistic traits, making them easier to learn by language models, regardless of the variation in the aspects of the service discussed.

The relatively high performance is also indicated by the negative sentiment class, with an F1 score of 0.88, a precision of 0.85, and a recall of 0.91. The higher recall value than the precision indicates that the model is quite sensitive in capturing patient dissatisfaction expressions, although there are still a number of misclassified negative predictions. In the context of healthcare communication, this suggests that patient complaints or criticisms are generally conveyed with fairly clear language markers, although not always uniformly.

In contrast, neutral sentiment showed the lowest performance among the three classes, with a precision of 0.66, a recall of 0.72, and an F1 score of 0.69. This value reflects the model's difficulty in distinguishing neutral expressions from positive and negative sentiments. Neutral expressions in patient feedback are often ambiguous, moderate, or implicit, so they have semantic proximity to the other two poles of sentiment. As a result, models tend to misclassify neutral expressions, a pattern consistent with the communication characteristics of patients who tend to avoid extreme statements.

Overall, the heat map confirms that the computational approach used is more effective in identifying polarized sentiments than ambiguous ones. These findings not only reflect the technical limitations of the model but also highlight the complexity of patient communication expressions in healthcare contexts, particularly on the use of neutral sentiment as an indirect communication strategy.

3.2. SROC Curve Summary

Figure 3 presents the ROC curve for each sentiment label as an indicator of the model's discriminating ability to distinguish sentiment classes. The AUC values obtained for positive (0.98) and negative (0.97) sentiments indicate that the model has an excellent ability to distinguish polarized sentiment expressions. These results indicate that the resulting feature representation can capture relatively consistent linguistic patterns in the expression of patient satisfaction and dissatisfaction.

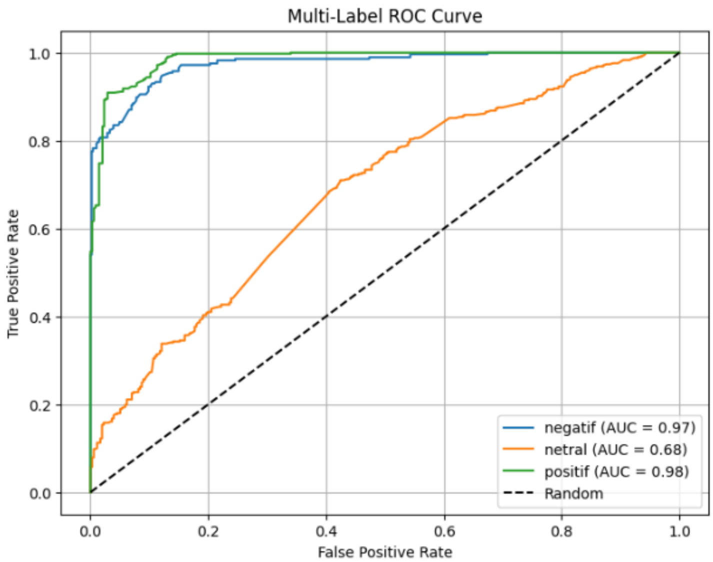

Figure 3. ROC-AUC scores per class indicate excellent model separability for polarized sentiments.

Conversely, a lower AUC value on neutral sentiment (0.68) indicates that the separation between neutral classes and other sentiment classes is still less than optimal. These findings indicate an overlap of feature characteristics between neutral expressions and positive and negative sentiments. In the context of patient feedback, it reflects the nature of neutral sentiments that are often conveyed implicitly or moderately, making it difficult to distinguish them explicitly based on linguistic signals alone.

Overall, the results of the ROC analysis show that the approach used is effective in distinguishing between clear and polarized sentiments, but still has limitations in identifying neutral sentiment expressions. These findings are positioned as an illustration of the characteristics of the data and the limits of the model's capabilities, as well as a basis for consideration for the development of a more sensitive sentiment analysis approach to moderate expression in future studies.

4. Discussion

The results of this study show that multilabel sentiment analysis based on a computational approach can provide a more structured understanding of patient communication patterns in healthcare feedback. The use of SMOTE proved particularly beneficial in balancing the training dataset, allowing the model to generalize better across minority sentiment classes. This finding aligns with previous work by Cahya et al. (2024), who emphasized the importance of oversampling techniques in boosting the performance of IndoBERT on imbalanced datasets.

The differences in performance between sentiment classes reflected in the evaluation results indicate that the patient's communication expressions are not homogeneous but have different levels of clarity and intensity depending on the type of sentiment conveyed. Higher performance on both positive and negative sentiment suggests that expressions of patient satisfaction and dissatisfaction are generally conveyed more explicitly in written feedback. In the context of healthcare communication, this indicates that patients tend to convey extreme judgments directly, both to appreciate the service they perceive to be satisfactory and to express complaints about disappointing experiences. This pattern allows a computational analysis approach to capture polarized messages with a relatively good degree of accuracy.

However, we encountered challenges in accurately classifying neutral sentiments, which obtained an F1-score of 0.69. This is consistent with findings by Aryanti et al. (2025), where IndoBERT struggled to detect neutral tones in mental health feedback. Likely due to the subtle linguistic markers that distinguish neutrality from polarity. Similarly, Yulianti & Nissa (2024) reported low precision in neutral sentiment detection when applying IndoBERT in ABSA tasks. The difficulty in neutral classification is further confirmed by the lower AUC value (0.68), indicating overlapping features and insufficient distinction from other classes.

The lower performance on neutral sentiment reflects the complexity of patient communication in conveying judgments that are moderate or ambivalent. Neutral sentiment expressions are often used as a non-confrontational communication strategy, in which patients convey implicit uncertainty, doubt, or dissatisfaction without using language that is clearly negative. This condition leads to an overlap of linguistic characteristics between neutral sentiment and polarized sentiment, making it difficult to explicitly separate classes in the classification process.

These findings suggest that the limitations in detecting neutral sentiment not only reflect the weaknesses of the technical approach but also reflect the nature of the patient's own communication. In healthcare communication practices, neutral sentiment can be an early indicator of fragile satisfaction or potential grievances that have not been explicitly articulated. Therefore, the results of neutral sentiment analysis need to be interpreted carefully so as not to be mistaken for the absence of problems in the service.

From the perspective of practical utilization, the computational approach used in this study can serve as a tool to systematically map patient communication tendencies, especially in identifying service areas that are often associated with positive and negative sentiments. This aligns with the research of Alkhnbashi et al. (2024), Mei et al. (2024), and Ihtada et al. (2025) Who also used IndoBERT and ABSA to support the implementation of a cross-domain model. Nevertheless, the results of the analysis are not intended to replace qualitative interpretation or professional judgment, but rather to complement the process of evaluating service communication based on patient feedback.

Overall, this discussion confirms that multilabel sentiment analysis based on computational approaches has the potential to support the evaluation of healthcare communication, while highlighting the importance of contextual understanding of patient sentiment expression. This approach is most effective when used as a supporting instrument that helps hospitals read patient communication patterns more systematically, without ignoring the complexity of meaning contained in the feedback narrative.

Table 3. Summarize previous research

Study	Domain	Task Type	SMOTE	F1-Score	Notes
Perwira et al. (2025)	Hospital Feedback	Multilabel	Yes	0.83	Balanced each aspect-sentiment via SMOTE
Aryanti et al. (2025).	Mental Health App	Single label	No	~0.79	Focus on binary classification
Ihtada et al. (2025).	E-Commerce	Multilabel	No	0.84	No resampling strategy applied
Mei et al. (2024).	Cosmetics Reviews	Multilabel	No	0.82	Lower recall for neutral classes
Alkhnbashi et al. (2024).	Clinical (Arabic/Eng)	Single label	Augment	0.81	Used LLM-based Arabic sentiment
Jazuli et al. (2023).	Student Feedback	ABSA	No	~0.80	Limited aspect granularity
Imaduddin et al. (2023).	Health App	Single label	No	0.78	Used IndoBERT without a multilabel setup

Although the performance of the approach used in this study is in line with the findings of previous studies, some limitations need to be acknowledged. First, the application of SMOTE as a data balancing strategy has the potential to produce synthetic samples that do not fully represent the linguistic complexity of natural texts. This condition can affect the sensitivity of the model, especially in distinguishing between neutral and contextual expressions of sentiment. In addition, the model architecture in this study relies on pre-defined service aspects, making it less flexible to handle feedback with dynamic or evolving aspect structures. Similar limitations are also noted by Perwira et al. (2025), who highlight the challenges of adapting domain-based models when applied to the context of changing feedback.

Another limitation relates to the data annotation process. Although agreement between annotators is maintained, the determination of sentiment labels, especially in the neutral category, still contains an element of subjectivity. This subjectivity has the potential to cause label noise that can affect the modeling learning process. Previous research has emphasized the importance of incremental annotation strategies and additional calibration procedures to minimize the impact of such noise in aspect-based sentiment analysis (Rani & Jain, 2023; Souza et al., 2023).

Despite these limitations, this research has relevant practical implications in the context of healthcare communication. The proposed approach can serve as a scalable analytical tool to map patient sentiment tendencies at the service aspect level. This information has the potential to assist hospital managers in identifying service areas that require further attention, such as facilities or support services, as part of an evaluation of patient feedback-based service communications. In addition, the use of data balancing strategies in the framework of multilabel classification can be an initial reference for future research that faces the problem of unbalanced label distribution in Indonesian textual data (Hadwan et al., 2022; Setiawan, 2024).

5. Conclusion

This study shows that multilabel sentiment analysis based on a computational approach can be used to read patient communication patterns in healthcare feedback in a more structured manner. The multilabel approach allows for a representation of sentiment towards different aspects of the service in a single feedback narrative, making it more appropriate to the multidimensional character of patient communication.

The findings of this study indicate that sentiment distribution imbalances need to be considered in the analysis of patient feedback, although the balancing strategies used are not intended as a universal solution. The use of the Indonesian pre-trained language model shows the potential for applying a computational approach to informal feedback data, with the note that the interpretation of the results is carried out contextually. The results also confirm that imbalanced sentiment distributions are an important factor to consider in patient feedback analysis. The application of a data balancing strategy, in this case per-label SMOTE, helps mitigate bias toward the majority class, but is not intended to be a universal solution or guarantee a fully balanced feature representation.

However, this study is limited by its data coverage from a single institution and the use of a predefined aspect scheme. Furthermore, it remains difficult to consistently capture neutral sentiment expressions. Further research is recommended to include cross-institutional data, develop a more dynamic aspect approach, and explore methods that are more sensitive to the ambiguity of patient language. Further research is recommended to: (1) conduct cross-institutional validation to test the robustness of the model across different service contexts; (2) explore more context-sensitive approaches for modeling neutral sentiment, such as contextual contrast learning or discourse-based modeling; (3) compare alternative balancing strategies that are more appropriate for text data; and (4) improve transparency and reproducibility through more detailed documentation of the methodology and the availability of supporting code or data. With a more careful and reflective approach, this study is expected to provide a starting point for the development of more mature multi-label sentiment analysis in the context of healthcare communication.

Reference

Alkhnbashi, O. S., Mohammad, R., & Hammoudeh, M. (2024). Aspect-Based Sentiment Analysis of Patient Feedback Using Large Language Models. *Big Data and Cognitive Computing*, *8*(12), 167. https://doi.org/10.3390/bdcc8120167

Aryanti, F. A. D., Luthfiarta, A., & Soeroso, D. A. I. (2025). Aspect-Based Sentiment Analysis with LDA and IndoBERT Algorithm on Mental Health App: Riliv. *Journal of Applied Informatics and Computing*, *9*(2), 361–375. https://doi.org/10.30871/jaic.v9i2.8958

Cahya, L. D., Luthfiarta, A., Krisna, J. I. T., Winarno, S., & Nugraha, A. (2024). Improving Multi-label Classification Performance on Imbalanced Datasets Through SMOTE Technique and Data Augmentation Using IndoBERT Model. *Jurnal Nasional Teknologi Dan Sistem Informasi*, *9*(3), 290–298. https://doi.org/10.25077/TEKNOSI.v9i3.2023.290-298

Hadwan, M., Al-Sarem, M., Saeed, F., & Al-Hagery, M. A. (2022). An Improved Sentiment Classification Approach for Measuring User Satisfaction toward Governmental Services' Mobile Apps Using Machine Learning Methods with Feature Engineering and SMOTE

Technique. *Applied Sciences*, *12*(11), 5547. https://doi.org/10.3390/app12115547

Ihtada, F. K., Alfianita, R., & Aziz, O. Q. (2025). Aspect-based Multilabel Classification of E-commerce Reviews Using Fine-tuned IndoBERT. *Kinetik: Game Technology, Information Systems, Computer Networks, Computing, Electronics, and Control*. https://doi.org/10.22219/kinetik.v10i1.2088

Imaduddin, H., A'la, F. Y., & Nugroho, Y. S. (2023). Sentiment Analysis in Indonesian Healthcare Applications using IndoBERT Approach. *International Journal of Advanced Computer Science and Applications*, *14*(8). https://doi.org/10.14569/IJACSA.2023.0140813

Jazuli, A., Widowati, & Kusumaningrum, R. (2023). Aspect-based sentiment analysis on student reviews using the Indo-Bert base model. *E3S Web of Conferences*, *448*, 02004. https://doi.org/10.1051/e3sconf/202344802004

Mei, N. C., Tiun, S., & Sastria, G. (2024). Multi-Label Aspect-Sentiment Classification on Indonesian Cosmetic Product Reviews with IndoBERT Model. *International Journal of Advanced Computer Science and Applications*, *15*(11). https://doi.org/10.14569/IJACSA.2024.0151168

Nazir, A., & Rao, Y. (2022). IAOTP. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1588–1598. https://doi.org/10.1145/3477495.3532085

Perwira, R. I., Permadi, V. A., Purnamasari, D. I., & Agusdin, R. P. (2025). Domain-Specific Fine-Tuning of IndoBERT for Aspect-Based Sentiment Analysis in Indonesian Travel User-Generated Content. *Journal of Information Systems Engineering and Business Intelligence*, *11*(1), 30–40. https://doi.org/10.20473/jisebi.11.1.30-40

Phan, M. H., & Ogunbona, P. O. (2020). Modelling Context and Syntactical Features for Aspect-based Sentiment Analysis. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 3211–3220. https://doi.org/10.18653/v1/2020.acl-main.293

Rani, S., & Jain, A. (2023). Aspect-based sentiment analysis of drug reviews using multi-task learning based dual BiLSTM model. *Multimedia Tools and Applications*, *83*(8), 22473–22501. https://doi.org/10.1007/s11042-023-16360-3

Rohman, S., & Agung, I. W. P. (2025). Komparasi algoritma indobert, svm, dan random forest dalam analisis sentimen program makan bergizi gratis. *Jurnal Mahasiswa Teknik Informatika*, *9*(6), 9464–9471. https://doi.org/https://doi.org/10.36040/jati.v9i6.15611

Setiawan, E. I. (2024). Aspect-Based Sentiment Analysis of Healthcare Reviews from Indonesian Hospitals based on Weighted Average Ensemble. *Journal of Applied Data Sciences*, *5*(4), 1579–1596. https://doi.org/10.47738/jads.v5i4.328

Souza, F. C., Nogueira, R. F., & Lotufo, R. A. (2023). BERT models for Brazilian Portuguese: Pretraining, evaluation and tokenization analysis. *Applied Soft Computing*, *149*, 110901. https://doi.org/10.1016/j.asoc.2023.110901

Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2020). IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding. *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, 843–857. https://doi.org/10.18653/v1/2020.aacl-main.85

Wong, E., Mavondo, F., & Fisher, J. (2020). Patient feedback to improve quality of patient-centred care in public hospitals: a systematic review of the evidence. *BMC Health Services Research*, *20*(1), 530. https://doi.org/10.1186/s12913-020-05383-3

Yulianti, E., & Nissa, N. K. (2024). ABSA of Indonesian customer reviews using IndoBERT: single- sentence and sentence-pair classification approaches. *Bulletin of Electrical Engineering and Informatics*, *13*(5), 3579–3589. https://doi.org/10.11591/eei.v13i5.8032