JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 31 No. 5, pp. 75-84, May 2026
https://doi.org/10.9708/jksci.2026.31.05.075

Optimizing Emotion Recognition in Korean Counseling Texts Using Parameter Efficient Fine Tuning

Myung-Suk Lee*

*Professor, Tabula Rasa College, Keimyung University, Deagu, Korea

[Abstract]

This study proposes a parameter-efficient emotion recognition model optimized for Korean counseling dialogues to address the demand for online psychological crisis response. Parameter-efficient fine-tuning (PEFT) techniques, specifically LoRA and QLoRA, were applied to KcBERT and evaluated against full fine-tuning and KoGPT2 using the AI Hub(Artificial Intelligence Hub) corpus. Results show that LoRA-based KcBERT achieved 0.824 accuracy and 0.812 F1-score while updating only 0.3% of parameters. t-SNE analysis confirmed stable emotional representation learning. This research demonstrates that high-performance emotion recognition can be practically deployed in resource-constrained environments, significantly improving the efficiency of AI-based mental health support.

▸ **Key words**: KcBERT, Korean counseling data, LoRA, Parameter-efficient fine tuning, Emotion recognition

[요　약]

본 연구는 온라인 심리 위기 대응 수요 증가에 대응하기 위해 한국어 상담 대화에 최적화된 파라미터 효율적 감정 인식 모델을 제안한다. 이를 위해 파라미터 효율적 미세조정(PEFT) 기법인 LoRA와 QLoRA를 KcBERT에 적용하고, AI Hub 코퍼스를 활용하여 전체 미세조정 및 KoGPT2와 성능을 비교·평가하였다. 실험 결과, LoRA를 적용한 KcBERT는 전체 파라미터의 0.3%만을 업데이트하면서도 정확도 0.824, F1-score 0.812를 달성하였다. 또한 t-SNE 분석을 통해 안정적인 감정 표현 학습이 이루어졌음을 확인하였다. 본 연구는 제한된 자원 환경에서도 고성능 감정 인식 모델의 실질적 배포 가능성을 입증하며, AI 기반 정신건강 지원 시스템의 효율성을 크게 향상시킬 수 있음을 보여준다.

▸ **주제어**: KcBERT, 한국어 상담 데이터, LoRA, 매개변수 효율 미세조정, 감정 인식

• First Author: Myung-Suk Lee, Corresponding Author: Myung-Suk Lee
*Myung-Suk Lee (mslee@kmu.ac.kr), Tabula Rasa College, Keimyung University
• Received: 2026. 04. 29, Revised: 2026. 05. 18, Accepted: 2026. 05. 19.

I. Introduction

Modern society has experienced a rapid migration of everyday activities toward online platforms as digital transformation accelerates, resulting in an unprecedented increase in individuals' online engagement. However, alongside this technological expansion, large volumes of conversational text data have been generated through online communication environments[1]. As text-based, non-face-to-face communication becomes increasingly common, analyzing such conversational data has emerged as an important challenge for understanding underlying emotional information[2].

At the core of emotion recognition from conversational text lies the ability to accurately capture subtle emotional cues embedded in linguistic expressions[3]. Recent advances in large language models(LLMs) have led to substantial improvements in natural language understanding (NLU) capabilities. Nevertheless, fine-tuning models with billions of parameters for domain-specific tasks remains a major challenge, as it demands extensive computational resources and significant training time.

Recent advancements in 2024 and 2025 have further highlighted the necessity of balancing model performance and computational overhead, particularly in sensitive domains like psychological counseling [4]. Emerging PEFT strategies emphasize not only parameter reduction but also the preservation of pre-trained knowledge during domain-specific adaptation [5].

This challenge is particularly pronounced for the Korean language, which is characterized as an agglutinative language with complex grammatical structures and frequent semantic ambiguity. Consequently, analyzing informal and unstructured conversational data requires optimization strategies that adequately reflect Korean-specific contextual nuances[6]. While conventional full fine-tuning approaches update all model parameters and can achieve high performance, they suffer from critical limitations in terms of computational efficiency. As an alternative, parameter-efficient fine-tuning (PEFT) techniques[7] - most notably Low-Rank Adaptation (LoRA)—have recently gained increasing attention.

Recent studies have explored the application of large language models and efficient fine-tuning strategies for emotion analysis and conversational understanding. For instance, instruction-tuned LLMs have demonstrated strong performance in emotion-cause analysis in conversational settings, highlighting the potential of LLM-based approaches for affective computing. In addition, efficient fine-tuning methods have been shown to significantly reduce computational cost while maintaining competitive performance in natural language understanding tasks[8].

The objective of this study is to conduct a comparative analysis of emotion recognition performance by applying PEFT techniques to both KcBERT, a representative Korean language-specific encoder model, and KoGPT2, a generative language model, using a large-scale Korean conversational emotion dataset. Specifically, this study systematically examines the effects of LoRA and QLoRA on classification accuracy and training efficiency across different model architectures. Through this investigation, we aim to empirically demonstrate that parameter-efficient fine-tuning methods can efficiently process emotional information in Korean conversational texts under resource-constrained environments.

While this study does not propose a new model architecture, its primary contribution lies in providing a systematic empirical analysis of PEFT methods in Korean counseling dialogue, a domain that has been relatively underexplored. In particular, this work highlights how parameter-efficient fine-tuning interacts with both encoder- and decoder-based architectures under realistic conversational settings.

II. Related work

1. KcBERT and KoGPT2

In the field of natural language processing (NLP), Transformer-based architectures are generally categorized into encoder-centric models and decoder-centric models according to their contextual modeling strategies. KcBERT (Korean comments BERT), employed in this study, is an encoder-based model specifically designed to capture bidirectional contextual relationships among words within a sentence[9]. By jointly considering both left and right contexts, encoder-based models are particularly effective for semantic understanding and text classification tasks.

KcBERT is pre-trained on large-scale Korean colloquial text, enabling robust handling of neologisms and non-standard expressions frequently observed in informal conversational data. In contrast, KoGPT2 is a decoder-based model that follows an auto-regressive learning paradigm, generating tokens sequentially based on previously generated context. While this architecture is well suited for text generation and continuation tasks, it may exhibit limitations in capturing holistic sentence-level emotional semantics compared to encoder-based models[10].

Accordingly, this study investigates how architectural differences between encoder- and decoder-based models influence emotion classification performance in Korean conversational text data.

2. Parameter-Efficient fine tuning

As the number of parameters in large-scale language models continues to increase, conventional full fine-tuning approaches—which require updating all model parameters—have become increasingly impractical due to substantial computational and memory demands. To address this limitation, parameter-efficient fine-tuning (PEFT) techniques have been proposed to improve training efficiency by freezing most pre-trained model parameters while updating only a small subset of task-specific weights.

Among various PEFT methods, Low-Rank Adaptation (LoRA) and Quantized LoRA (QLoRA) have attracted considerable attention. LoRA decomposes the task-specific weight update into two low-rank matrices, enabling effective adaptation with a significantly reduced number of trainable parameters[11]. Let ($W_0 \in R^{d \times k}$) denote the pre-trained weight matrix. Instead of directly updating, LoRA represents the task-specific weight update ΔW as a low-rank decomposition, as expressed in Equation (1).

$$h = W_{0x} + \Delta W_x = W_{0x} + BA_x \quad \text{Equation (1)}$$

As illustrated in Equation (1), during the forward pass, the input representation is processed simultaneously through the frozen pre-trained weight matrix and the trainable low-rank matrices, and their outputs are combined to produce the final output. This design allows only a very small fraction of the total parameters—approximately 0.3% in this study—to be updated, thereby substantially reducing computational overhead while maintaining performance comparable to that of full fine-tuning.

QLoRA further extends this approach by incorporating low-bit quantization to reduce memory consumption without significantly degrading model performance. By applying 4-bit quantization to base model parameters, QLoRA achieves substantial memory efficiency while retaining effective fine-tuning capability. These techniques enable large-scale language models to be adapted efficiently to domain-specific tasks under resource-constrained environments[12].

Figure 1 illustrates the structural differences among full fine-tuning, LoRA, and QLoRA, highlighting the contrast between conventional approaches that update all model parameters and PEFT-based methods that selectively fine-tune a limited subset of weights[12].

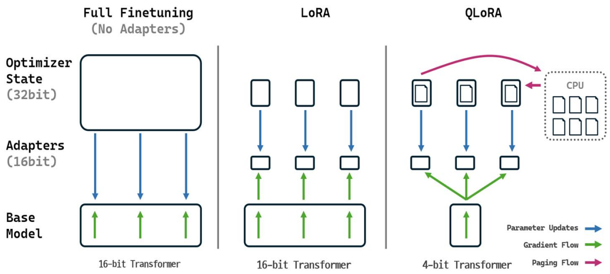

Fig. 1. Comparison of Architectures by Fine-tuning Techniques[12]

To identify suitable configurations for Korean emotion recognition, the structural characteristics and training specifications of each model and fine-tuning strategy are summarized in Table 1.

Table 1. Comparison of Technical Specifications by Experimental Model and Fine-tuning Method

Category	KoGPT2 (Full FT)	LoRA (KcBERT)	QLoRA (KcBERT)
Model Structure	Decoder(Optimized for Generation)	Encoder (Optimized for Understanding)	Encoder (Optimized for Understanding)
Trainable Parameters	Full (100%)	Partial (approx. 0.3%)	Partial (approx. 0.3% + Quantization)
Computing Resources	Very High	Low	Very Low (4-bit)
Key Advantages	Sentence Generation & Scalability	Fast Training Speed & High Performance	Extreme Memory Efficiency

Recent studies in 2024-2025 explored parameter-efficient fine-tuning and emotion-aware conversational AI systems. Kim and Kim (2024)[13] investigated LoRA-based fine-tuning using curriculum learning in resource-constrained environments, while Ji (2025)[14] proposed an emotion-aware conversational model integrating RAG-based response generation. However, these studies mainly focused on document classification or conversational response generation tasks.

In contrast, this study specifically targets Korean counseling dialogue data and comparatively analyzes LoRA- and QLoRA-based fine-tuning approaches for multi-class emotion classification. Furthermore, this work jointly considers classification performance, computational efficiency, confusion matrix analysis, t-SNE visualization, and practical sentence-level emotion prediction within counseling-oriented scenarios.

III. Methodology

1. Dataset and preprocessing

This study utilized the Empathy-based Dialogue Corpusprovided by AI Hub(Artificial Intelligence Hub)[15]. The dataset contains multi-layered emotional annotations derived from real-world counseling scenarios. For the purpose of this research, six major emotional categories were selected: Joy, Sadness, Anger, Anxiety, Hurt, and Embarrassment. The dataset comprises a total of 51,630 records. The dataset distribution across the six emotion categories was analyzed prior to training. As shown in Table 2, moderate class imbalance was observed among emotion labels, particularly in the Joy category, which contained fewer samples compared to the other classes.

Table 2. Comparison of Technical Specifications by Experimental Model and Fine-tuning Method

Emotion Category	Number of Samples	percentage
Anger	9,160	17.7%
Anxiety	9,320	18.1%
Hurt	9,143	17.7%
Embarrassment	8,756	16.9%
Sadness	9,125	17.7%
Joy	6,126	11.9%
Total	51,630	100%

The overall structure illustrates relationships among age, gender, situational keywords, and conversational context (Human Sentence 1-3). Data preprocessing applied text cleansing to remove special characters and redundant whitespaces. To preserve contextual characteristics, Human Sentence 1-3 were concatenated into a single input representing the user's emotional flow across the dialogue.

In this study, only user utterances were used, while counselor responses were excluded to focus on modeling the speaker's emotional trajectory.

Counselor responses mainly consisted of neutral counseling expressions such as empathy, guidance, and follow-up questions. Although excluding counselor responses may limit contextual completeness, sequential user utterances were considered sufficient to capture emotional progression for classification.

Tokenization was conducted using the dedicated tokenizers of KcBERT and KoGPT2, converting text into integer indices. Finally, the dataset was divided into training, validation, and test sets with an 8:1:1 ratio to ensure robustness and objectivity.

2. Experimental environment and LoRA fine tuning configuration

All experiments were conducted in an NVIDIA T4 GPU environment using Python and the PyTorch framework. The key hyperparameter settings for optimizing the LoRA-based emotion recognition models are summarized in Table 3.

To ensure experimental consistency and fair comparison, all LoRA-based experiments, including those for KcBERT, were re-conducted under identical training conditions after incorporating KoGPT2 with LoRA.

Table 3. Key Hyperparameter Settings for LoRA and Training

Parameter	Value	Notes
Rank (γ)	8	Size of the low-rank matrices
LoRA Alpha (α)	16	Weight update scaling
Target Modules	Query, Value	Applied consistently to both KcBERT and KoGPT2 for fair comparison
Learning Rate	2e-4	Learning rate
Batch Size	32	Batch size
Epochs	3	Number of training epochs
Precision	FP16 / 4-bit	4-bit used when applying QLoRA

IV. Results

This section presents a comprehensive evaluation of emotion classification performance under different model architectures and fine-tuning strategies. Quantitative performance metrics are first compared across models, followed by efficiency analysis of LoRA-based fine-tuning. Finally, qualitative verification is conducted using embedding visualization and confusion matrix analysis to assess semantic separability and classification reliability.

1. Performance comparison across models and training strategies

To evaluate the effectiveness of parameter-efficient fine-tuning, emotion classification performance was compared across two model architectures—KoGPT2 and KcBERT—and three training strategies: Full Fine-Tuning, LoRA, and QLoRA. Performance was assessed using Accuracy, Precision, Recall, and F1 Score, as summarized in Table 4. To ensure a fair comparison of parameter-efficient fine-tuning across architectures, LoRA was additionally applied to KoGPT2, allowing direct evaluation of PEFT effects within both encoder- and decoder-based models.

Table 4. Performance Metrics Comparison by Model and Training Technique

Model & Method	Accuracy	Precision	Recall	F1 Score
KoGPT2 (Full FT)	0.762	0.758	0.741	0.749
KoGPT2 (LoRA)	0.822	0.805	0.816	0.808
KcBERT(Full FT)	0.815	0.803	0.808	0.805
KcBERT (LoRA)	0.824	0.815	0.809	0.812
KcBERT (QLoRA)	0.811	0.802	0.795	0.798

The results indicate that KcBERT consistently outperformed KoGPT2 across all evaluation metrics, confirming that an encoder-based architecture is more suitable for capturing contextual information in Korean counseling dialogues. Among the optimization techniques, the KcBERT-LoRA model achieved the highest performance, with an accuracy of 0.824 and an F1 score of 0.812. Although QLoRA demonstrated competitive results, its performance was slightly lower than LoRA, which can be attributed to information loss introduced by aggressive quantization.

The superior performance of LoRA compared to full fine-tuning can be attributed to its implicit regularization effect. Full fine-tuning updates all model parameters, which can lead to overfitting, particularly in domain-specific datasets with limited diversity. In contrast, LoRA constrains parameter updates to low-rank subspaces, preserving pre-trained knowledge while enabling task-specific adaptation. This effect is more pronounced in KoGPT2, where LoRA significantly improves performance, suggesting that parameter-efficient adaptation plays a critical role for decoder-based models that are not inherently optimized for classification tasks.

In addition to performance improvements, LoRA significantly reduces computational cost. Specifically, the number of trainable parameters was reduced to approximately 0.3% compared to full fine-tuning. Furthermore, GPU memory usage and training time were reduced by approximately 50% and 40%, respectively, demonstrating the practical efficiency of PEFT methods.

These findings suggest that LoRA provides an effective balance between performance and computational efficiency, making it preferable to both full fine-tuning and QLoRA for precise emotion classification tasks.

2. Efficiency analysis of LoRA-based KcBERT

In addition to classification performance, training efficiency and stability were analyzed for the LoRA-based KcBERT model. Training and validation loss values were monitored across epochs to assess convergence behavior.The results are presented in Table 5.

Table 5. Training and Validation Loss Trends per Epoch

Epoch	Training Loss	Validation Loss	Status
1	1.047000	1,066606	Training Started
2	1.015400	1.024295	Improving
3	0.964100	1.013501	Optimized

Both training and validation losses showed a steady decline over the training epochs, indicating stable learning without signs of overfitting. The validation loss reached its minimum at the final epoch, suggesting that LoRA effectively adapts the pre-trained model to the counseling domain while preserving generalization capability.

The overall fine-tuning procedure using LoRA is summarized in Table 6, which outlines the key stages from data preprocessing to final model generation.

Table 6. Training and Validation Loss Trends per Epoch

Phase	Key Components	Description
1. Data Preparation	Data loading, labeling, context construction	Counseling dialogue data were loaded from CSV files, and emotion labels (e.g., Joy, Sadness) were converted into numerical indices. Three dialogue turns were concatenated to preserve conversational context, followed by a 90:10 split into training and validation sets.
2. LoRA-based Model Adaptation	Tokenization, LoRA configuration, training setup	Input texts were tokenized and standardized to a fixed length of 128 tokens. LoRA was applied by injecting trainable low-rank adapters into the Query and Value projection layers of both KcBERT and KOGPT2, enabling parameter efficient fine tuning across different model architectures.
3. Training and Model Optimization	Training execution, validation, model saving	Model training was conducted over three epochs with periodic validation. The best-performing checkpoint was selected based on validation loss, resulting in an optimized counselingoriented emotion classification model.

Notably, LoRA updated approximately 0.3% of the total model parameters, resulting in a substantial reduction in computational and memory requirements compared to full fine-tuning. These findings indicate relatively stable generalization performance within the validation environment.

3. Qualitative analysis using t-SNE and confusion matrix

To further verify the qualitative performance of the proposed KcBERT-LoRA model, qualitative analyses were conducted using unseen counseling sentences, embedding visualization, and confusion matrix evaluation. First, model inference performance was examined using randomly selected counseling sentences that were not included in the training data. The emotion prediction results and corresponding confidence scores are illustrated in Table 7.

Table 7. Results of emotion prediction and confidence score verification for untrained counseling sentences.

Sentence	Predicted Emotion (Confidence)
After talking with the counselor. I feel much lighter. Thank you.	Joy 89.14%
No matter how hard I try, the situation does not improve. And I feel very frustrated and angry.	Anger: 87.24%
I am not sure if I can do well. I keep feeling anxious.	Anxiety: 89.52%
This happened suddenly, and I am very confused and do not know what to do.	Embarrassment: 72.88%

The results show that the model accurately captured subtle emotional cues within contextualized counseling expressions. For example, expressions of gratitude were classified as Joy with high confidence, while sentences describing frustration and anger were correctly identified as Anger. Similarly, expressions indicating uncertainty and sudden situational change were successfully classified as Anxiety and Embarrassment, respectively.

To assess real-world applicability, user-based testing was conducted using a real-time emotion analysis interface, as illustrated in Figure 4. The system accepts user input, preprocesses the text, and feeds it into a pre-trained KcBERT model with LoRA applied. The model outputs both an emotion label and a corresponding prediction probability, which is interpreted as a confidence score.

The predicted emotion and its confidence score are presented to the user in percentage form. Additionally, when the confidence score falls below a predefined threshold (e.g., 0.5), the system provides a guidance message encouraging the user to clarify the input, considering potential ambiguity in the sentence. The interface is designed to support real-time interaction while prioritizing the interpretability of model outputs.

Fig. 2. Emotion analysis testing and real-time feedback interface based on user text input

The results indicate that the model performs particularly well on colloquial and context-rich sentences resembling real counseling conversations. Longer, context-rich inputs tend to yield higher confidence scores, whereas shorter or more ambiguous expressions results in relatively lower confidence, reflecting the inherent complexity of counseling language.

To analyze the internal semantic structure learned by the model, t-SNE visualization was applied to the embedding space. The resulting distribution is shown in Figure 5.

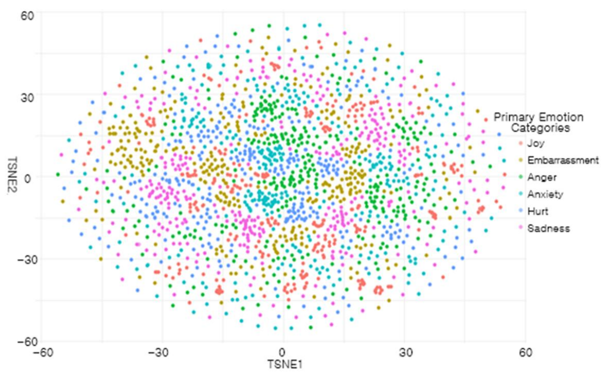

Fig. 3. Multidimensional Spatial Distribution and Cluster Analysis of Emotion Embeddings Using the t-SNE Algorithm

The visualization reveals a dense central cluster composed of Anger, Anxiety, and Hurt, suggesting strong semantic overlap among these emotions in counseling contexts. In contrast, Joy exhibited the widest dispersion, indicating a broad lexical variety associated with positive emotional expressions. These patterns demonstrate that the model captures realistic emotional relationships rather than relying solely on surface-level keywords[14].

Finally, classification reliability was assessed using a confusion matrix for the six emotional categories, as presented in Figure 6.

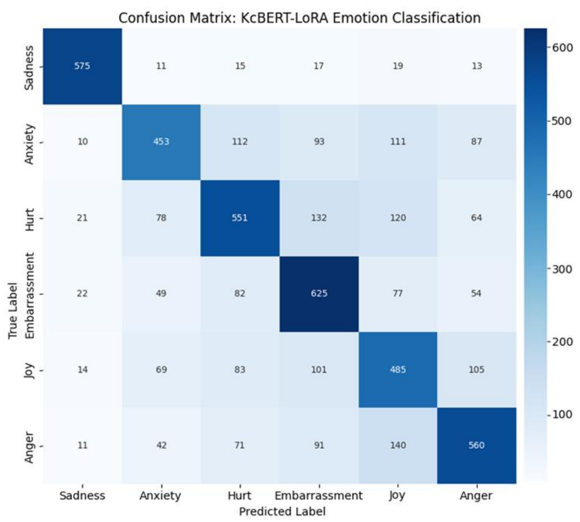

Fig. 4. Confusion matrix results of the KcBERT-LoRA-based emotion classification model

Strong diagonal dominance was observed across all categories, indicating high overall classification accuracy. In particular, Embarrassment, Sadness, and Anger showed high correct classification rates, while moderate confusion was observed among Anxiety, Hurt, and Joy. This confusion reflects the natural overlap of emotional expressions in counseling narratives and highlights the complexity of emotion recognition as an information-processing task.

This confusion reflects the natural overlap between semantically related emotional expressions in counseling conversations and suggests that future improvements may benefit from context-aware dialogue modeling or multi-label emotion classification approaches.

4. Class-wise emotion classification performance

To further evaluate model robustness under class imbalance conditions, class-wise precision, recall, F1-score, and support values were additionally analyzed for the KcBERT-LoRA model.

Table 8. Emotion Classification Performance

Emotion Category	Precision	Recall	F1-Score	Support
Anger	0.8777	0.9256	0.9010	605
Sadness	0.6769	0.6578	0.6326	979
Anxiety	0.6302	0.6032	0.6164	935
Hurt	0.6461	0.5836	0.6133	807
Embarrassment	0.5983	0.6079	0.6031	931
Joy	0.5622	0.5331	0.5355	906
Macro Avg	**0.6499**	**0.6519**	**0.6503**	**5163**
Weighted Avg	**0.6528**	**0.6368**	**0.6356**	**5163**
Accuracy	**-**	**-**	**0.6368**	**5163**

The experimental results indicate that the Anger category achieved the highest classification performance, whereas the Joy category showed relatively lower F1-scores. This tendency may be partially attributed to the moderate class imbalance observed in the dataset, where the Joy category contained fewer training samples compared to the other emotion classes. To provide a more balanced evaluation under class imbalance conditions, both Macro-F1 and Weighted-F1 metrics were additionally reported.

V. Conclusions

In this study, we proposed a Korean emotion recognition model based on Low-Rank Adaptation (LoRA), a representative Parameter-Efficient Fine-Tuning (PEFT) technique, as a technical approach to supporting non-face-to-face psychological counseling services. The practical applicability of the proposed model was further validated through its deployment in a mobile application environment.

From a performance perspective, applying LoRA to KcBERT, which is optimized for colloquial Korean text, achieved the best classification results, with an accuracy of 0.824 and an F1-score of 0.812. These findings indicate that encoder-based language models with bidirectional contextual modeling capabilities are more effective for capturing complex emotional states in counseling dialogues than decoder-based generative models.

In terms of computational efficiency, the results confirmed that LoRA and QLoRA can achieve performance comparable to full fine-tuning while updating only approximately 0.3% of the total model parameters. This substantial reduction in memory consumption and training cost provides a practical foundation for deploying emotion recognition systems in resource-constrained environments, such as small- and medium-scale counseling centers.

To evaluate practical scalability, the trained model was integrated into a mobile application, Mind Care AI. Real-time user testing demonstrated that the system can effectively capture subtle emotional changes in counseling dialogues and present the results in an interpretable visual format. This suggests that the proposed system can support counselors by providing objective emotional insights while also offering immediate feedback to users in online counseling settings.

Despite these contributions, certain limitations were observed, including prediction bias in specific emotion categories due to class imbalance in the dataset. Future work will address this issue through data augmentation and weighted loss functions. In addition, the current single-label classification framework will be extended to a multi-label setting to better reflect the coexistence of mixed emotions in real counseling conversations.

Furthermore, this study primarily focused on validating the effectiveness of LoRA-based emotion classification under resource-constrained environments. Comparative analysis with other parameter-efficient fine-tuning techniques, such as Adapter and Prefix-Tuning, was beyond the scope of the current study and remains an important direction for future research. This study provides a practical and scalable framework for applying parameter-efficient language models to intelligent mental health support systems.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2024S1A5A2A01019590)

REFERENCES

[1] L. Li, Y. Zhang, S. Li, and X. Chen, "Screen time and depression risk: A meta-analysis of cohort studies," Front. Psychiatry, 13, Article. 1058572, 2022. https://doi.org/10.3389/fpsyt.2022.1058572

[2] B. A. Primack et al., "Social media use and perceived social isolation among young adults in the U.S.," Am. J. Prev. Med., 53(1), pp. 1–8, 2017. https://doi.org/10.1016/j.amepre.2017.01.010

[3] B. Pang and L. Lee, "Opinion mining and sentiment analysis," Foundations and Trends in Information Retrieval, 2(1–2), pp. 1–135, 2008. http://dx.doi.org/10.1561/1500000001

[4] Y. Wu et al., "A comprehensive survey of forgetting in LLMs: Parameters, data, and fine-tuning," IEEE Transactions on Knowledge and Data Engineering, 2024 (Early Access) https://arxiv.org/html/2307.09218v3

[5] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners (arXiv:2005.14165). https://arxiv.org/abs/2005.14165

[6] S. Lee, H. Jang, Y. Baik, S. Park, and H. Shin, "KR-BERT: A small-scale Korean-specific language model," Journal of KIISE, 47(7), pp. 682–692, 2020. https://doi.org/10.5626/JOK.2020.47.7.682

[7] N. Ding et al., "Parameter-efficient fine-tuning of large-scale pre-trained language models," Nature Machine Intelligence, 5(3), pp. 220-235, 2023.

[8] Luo, M., Zhang, H., Wu, S., Li, B., Han, H., & Fei, H. (2024). NUS-Emo at SemEval-2024 Task 3: Instruction-tuning LLM for multimodal emotion-cause analysis in conversations. Proceedings of SemEval-2024.

[9] J. Lee, "KcBERT: Korean comments BERT," Proceedings of the 32nd Conference on Hangul and Korean Information Processing (HCLT), 2020. https://www.koreascience.or.kr/article/CFKO202030060691828.pdf

[10] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever,

"Language models are unsupervised multi-task learners," OpenAI Blog, 1(8), pp. 9, 2019. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf

[11] E. J. Hu et al., "LoRA: Low-rank adaptation of large language models," in Proc. Int. Conf. Learn. Represent. (ICLR), 2022 https://openreview.net/forum?id=nZeVKeeFYf9

[12] T. Dettmers, A. Pagnoni, A. Holtzman, and L. Zettlemoyer, "QLoRA: Efficient finetuning of quantized LLMs," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 36, pp. 10088-10115, 2023.

[13] D. Kim and N. Kim, “Enhancing LoRA Fine-tuning Performance Using Curriculum Learning,” Journal of The Korea Society of Computer and Information, vol. 29, no. 3, pp. 43?54, 2024.

[14] S. Ji, “Customer Emotion-Aware Conversational Language Model using Generative AI and Retrieval-Augmented Generation (RAG),” Journal of The Korea Society of Computer and Information, vol. 30, no. 8, pp. 103?110, 2025.

[15] AI Hub, "Empathy-based dialogue corpus," [Dataset], 2025. [Online]. Available: https://aihub.or.kr/aidata/7976 (Accessed: Nov. 2025).

[16] M. S. Lee, "A study on emotion patterns using the emotion analysis corpus dataset," in Proc. KSCI Conf., 33(2), 2025. https://ksci.re.kr/board/snotice/article/263483

Authors

Myung-Suk Lee received the M.S. and Ph.D. degrees in Computer Engineering from Keimyung University, Korea, in 2003 and 2009, respectively. Dr. Lee joined the faculty of Tabula Rasa College at Keimyung University, Daegu, Korea, in 2013. She is currently an Associate Professor in Tabula Rasa College, Keimyung University. Her research interests include artificial intelligence, learning analytics, computer education, and networks.