KoBERT-Based Multi-Class Text Classification Applied to *Act on Protective Action Guidelines Against Radiation in the Natural Environment*

Hyunwoo Lee[a, b], Jongeun Kim[a, c], Minjung Kim[a], Sujin Kim[a, b], Insu Chang[b], Seungkyu Lee[b], Yoonsun Chung[a*]

[a]Department of Nuclear Engineering, Hanyang university, 222, Wangsimni-ro, Seongdong-gu, Seoul, Korea

[b]Korea Atomic Energy Research Institute, 111 Daedeok-Daero, Yuseong-gu, Daejeon, Korea

[c]Proton Therapy Center, National Cancer Center, 323, Ilsan-ro, Ilsandong-gu, Goyang-si, Gyeonggi-do, Korea

*Corresponding author: ychung@hanyang.ac.kr

Introduction

- **Large Language Models(LLMs) trained on general-purpose text data** often exhibit **suboptimal performance in domain-specific contexts**.
  → The most common approach : **fine-tuning**.
- **The performance of language models significantly improves** after fine-tuning across various tasks, including **question answering** (Q&A), **named entity recognition** (NER), **summarization**, and **text classification**.
- In this study, we **pre-trained** **KoBERT**, which is a BERT-based Korean language model on domain-specific texts related to the ***Act on Protective Action Guidelines Against Radiation in the Natural Environment*** and **fine-tuned** it to improve performance on **text classification** in that domain.
- Finally, the **performance of the proposed domain-specific model was compared with that of commercial LLM** to assess whether a specialized language model could achieve performance comparable to general-purpose commercial systems.

Materials and Methods

BERT and KoBERT

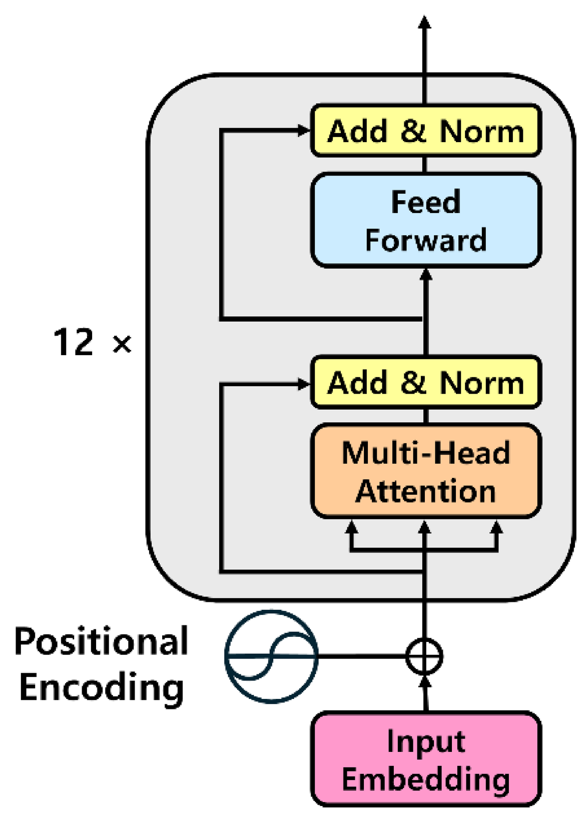

Figure 1. Structure of BERT

BERT

- **B**idirectional **E**ncoder **R**epresentations from **T**ransformers (BERT)
- Uses a **multi-head attention mechanism** to analyze contextual relationships among words

KoBERT

- **KoBERT** is a **lighter version of BERT** developed for the **Korean language**
- **Vocabulary** (30,002→8,002) and **parameters** (110M→92M) are reduced in KoBERT

Dataset and Vocabulary

Dataset for pre-training and fine-tuning was collected from four sources.

- Investigation and Analysis of Actual State of Safety Management for Radiation in the Natural Environment reports (2014–2022) [as "Reports"]
- Nuclear Safety Yearbook volumes (2016, 2017, 2020, 2022) [as "Yearbooks"]
- Internet news articles retrieved using the keyword "Radiation" [as "News"]
- Abstracts of research papers related to radiation in the natural environment [as "Abstracts"]

Table 1. Size of dataset from each data sources

Reports	Yearbooks	News	Abstracts
170 KB	179 KB	23,000 KB	72 KB

Pre-training & Fine-tuning

Domain-Adaptive Pre-Training (DAPT)

- To teach model **the linguistic characteristics** of the target domain
- **Masked Language Modeling** (MLM) and **Next Sentence Prediction** (NSP)
- 100 epochs, learning rate = 1e-04

Fine-tuning

- To improve **classification** performance of the model
- Uses **[CLS]** token for classification
- 10 epochs, learning rate = 1e-05
- The **model returns** **clause number** **of the *Act on Protective Action Guidelines Against Radiation in the Natural Environment*** or **"Not related"**, depends on the content and context of the input sentence.

Metrics

$$Total\ accuracy = \frac{Number\ of\ correct\ answers}{Number\ of\ samples\ in\ test\ set}$$

$$Precision = \sum_{i=1}^{N} \frac{P_i}{N}, \qquad P_i = precision\ of\ class\ i$$

$$Recall = \sum_{i=1}^{N} \frac{R_i}{N}, \qquad R_i = recall\ of\ class\ i$$

$$macro\ F1\ score = \sum_{i=1}^{N} \left(\frac{2 \times P_i \times R_i}{P_i + R_i} \right) / N$$

Results and Discussion

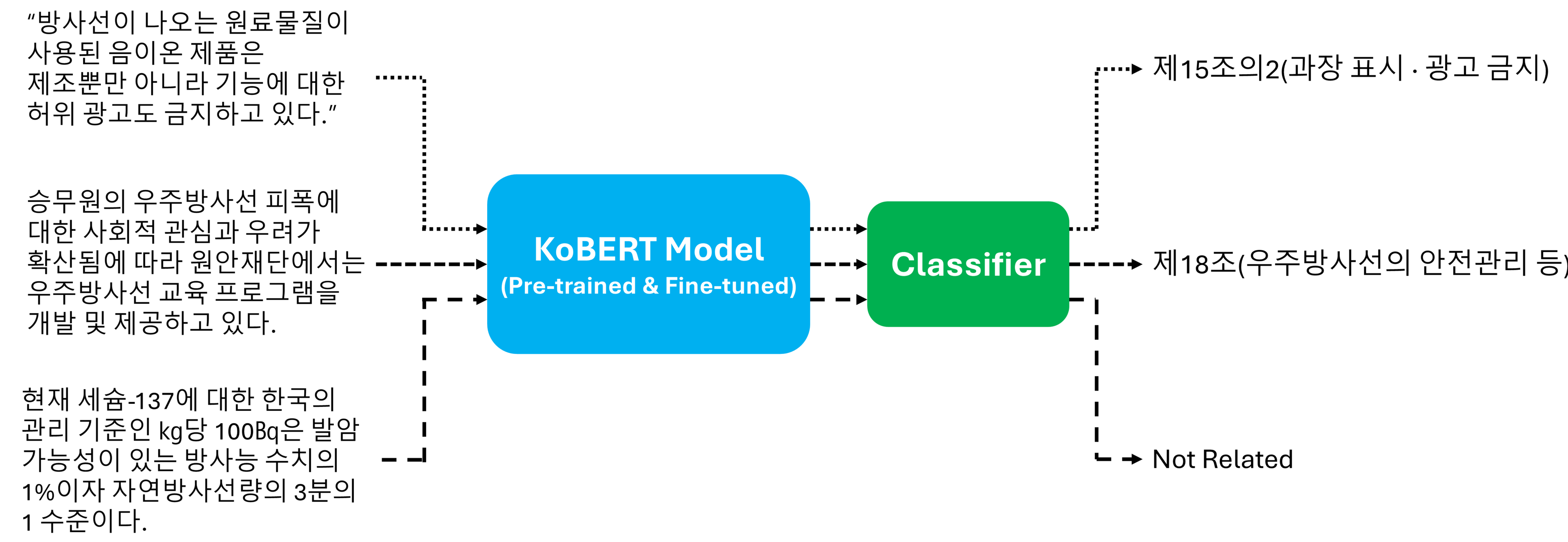

Figure 2. Schematic diagram and example of language model-based text classifier

Table 2. Precision, Recall, macro F1 score of models

Model	Total accuracy	Precision	Recall	Macro F1
KoBERT				
w/o DAPT, CE*	**0.88**	0.21	0.17	0.18
w/o DAPT, wCE**	0.83	0.35	0.41	0.35
w/ DAPT, CE*	**0.88**	0.15	0.18	0.16
w/ DAPT, wCE**	0.82	**0.36**	**0.45**	**0.37**
GPT-5-mini	0.82	0.23	0.09	0.12

*CE : Cross-Entropy loss function
**wCE : weighted Cross-Entropy loss function

Impact of loss function

- The **weighted cross-entropy loss function** proved highly effective in **mitigating the performance gap between major and minor classes**.
- However, it also led to a slight decrease in total accuracy, as it reduced the model's ability to correctly classify the majority "Not related" class.
- The **selection of appropriate loss function** **is equally(or sometimes even more) important** than conducting pre-training.

Impact of pre-training

- The case **KoBERT (DAPT, wCE)** showed an **improvement** in **macro precision, recall, and F1 score** compared to KoBERT (Base, wCE).
- This result demonstrate that **acquiring domain-specific knowledge** **is beneficial for complex multi-class classification tasks.**

Comparison with GPT-5-mini

- GPT-5-mini achieved **total accuracy comparable to the best KoBERT models**.
- However, due to very low recall, it resulted in the **lowest macro F1 score** among all tested models.
- This suggests that for complex, domain-specific tasks, **fine-tuning** **remains a more effective approach** than using a general-purpose commercial LLM.

Conclusion

- In this study, **Domain-Adaptive Pre-Training (DAPT)** and **fine-tuning** was performed on KoBERT using a **domain-specific corpus on radiation regulations**.
- **Four configurations of the fine-tuned KoBERT**—with and without DAPT and using either cross-entropy or weighted cross-entropy loss—was tested and their performance was compared to that of GPT-5-mini.
- In terms of total accuracy, the best-performing models were **KoBERT (Base, CE)** and **KoBERT (DAPT, CE)**, while **KoBERT (DAPT, wCE)** **achieved the highest macro F1 score.**
- **GPT-5-mini** achieved **total accuracy of 0.82**, which is equivalent to that of KoBERT(DAPT, wCE), but **lowest macro f1 score** among all models.
- These results suggest that **fine-tuning** **constitutes** **a competitive method** **within** **domain-specific contexts**, and that **other linguistic tasks related to nuclear energy and radiation could be addressed using the same strategy.**