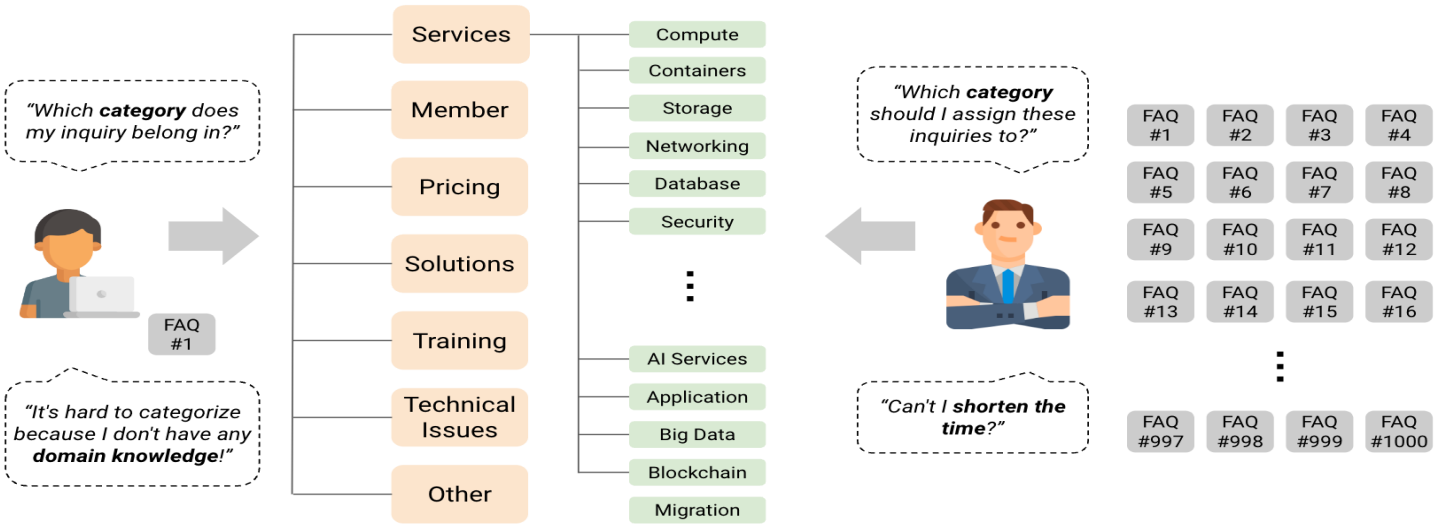
Improving Cloud FAQ Experience through Contrastive Learning-based Inquiry Classification*

Yongshin Kim○ Taehee Lee Sanghyeon Jung Chanjae Lee Taewan Kwon
OKESTRO
{ys.kim2, th.lee, sh.jung, cjlee, tw.kwon}@okestro.com

Abstract

Users constantly ask questions, and the company provides Frequently Asked Questions (FAQ) to make users write inquiries. This paper contributes to enhancing the experience of users and administrators in the FAQ system. This paper presents a model that uses BERT and contrastive learning to automatically classify users' inquiries. The model eliminates the need for users to have domain knowledge or for administrators to classify multiple inquiries, which can significantly improve the efficiency of the inquiry management system, reduce the workload of users and administrators, and enhance the overall user experience. The study collected 2,576 question-and-answer pairs of data from cloud platforms and used Korean comments BERT (KcBERT) pretrained language model to perform Inquiry classification. To more accurately predict the categories of user inquiries, this study trained the model using contrastive loss that enhances the similarity of questions and answers as well as classification loss.



<Figure 1> When the users leave an Inquiries, the FAQ system is divided into cases where the user must classify the category of the inquiry themselves and where the administrator must classify a large amount of user inquiries.

1. Introduction

As companies and organizations increasingly rely on online FAQ systems to manage user inquiries, it is essential to provide a seamless and user-friendly experience for customers. However, the current system presents some challenges for users with limited domain knowledge and a small number of administrators who need to classify multiple inquiries. <Figure 1> shows problems that may occur when registering inquiries in the cloud domain from a user's perspective and from a administrator's perspective. On the Naver cloud platform(https://www.ncloud.com/), users must directly select one of the 100 categories and leave their inquiries. This is a time-consuming and challenging task, especially for users who lack domain knowledge. In addition, if the user does not classify the inquiries correctly, they will not be passed to the right

*This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00147, Technology of Integrated Management Platform for Multi-Hybrid SaaS Solution)

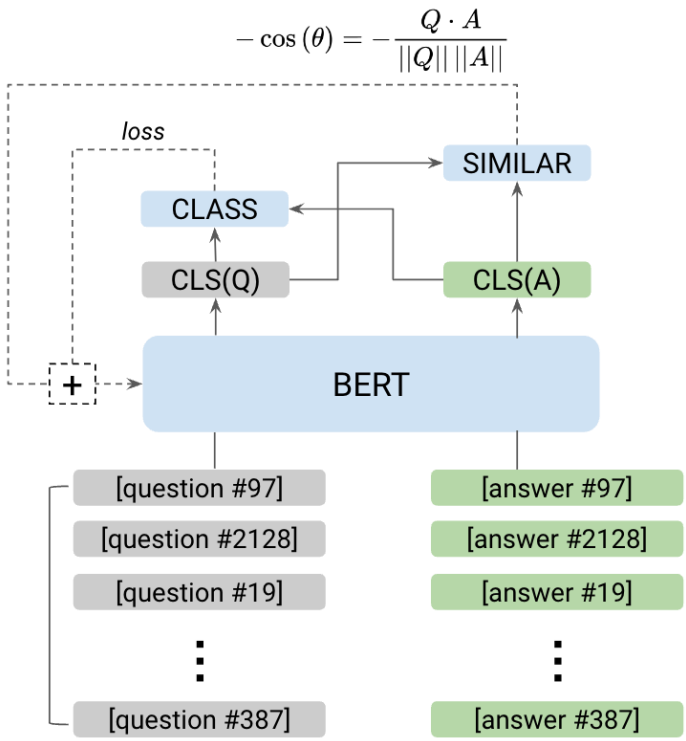
administrator, which can lead to delayed and inaccurate response.

On the other hand, in some cases, the user does not directly classify the category of the inquiry but simply submits it. For example, the user of kakaoicloud (https://www.kakaoicloud.com/) just writes and submits the inquiry. From the user's point of view, there seems to be no problem. However, the user's inquiries must be classified directly by the administrator and delivered to the person in charge. If a small number of administrators are given multiple inquiries, classification will take a considerable amount of time.

To address these challenges, this study presents a model that can manage users' inquiries using BERT (Bidirectional Encoder Representations from Transformers) [1], which is widely used in the field of natural language processing. This approach does not require users to have the domain knowledge to categorize their inquiries or administrators to classify multiple inquiries. Instead, the inquiries are automatically categorized into appropriate categories by a deep learning algorithm. It has the potential to significantly improve the efficiency of the inquiry management system, reduce the workload of users and administrators, and enhance the overall user experience.

Similar to our study, attempts to enhance the FAQ system have existed several times before. Shaw et al. [2] proposed an automatic FAQ classification system using ontology to improve user satisfaction by reducing inconsistencies that occur when administrators manually classify FAQ information and by reducing knowledge search time. Daman et al. [3] proposed a model called MMT-DNN that uses a multi-transformer to solve the difficulties users face in finding questions most similar to their own on the FAQ system. Mitra et a. [4] proposed a BERT-based model that recommends more accurate and relevant questions for user questions to address the negative user experience that arises when a search engine recommends unrelated questions. The above studies are similar to our studies in that they have improved the user experience of the FAQ system, but they have not considered the time cost of users and administrators that can be obtained by automatically classifying inquiries.

In this study, we present a model that achieves high accuracy in classifying inquiries while increasing the similarity between inquiries and answers using contrastive learning. We will evaluate the effectiveness of the proposed approach and discuss its implications for FAQ management. The contribution of this study is as follows. (1) We present a solution to the challenges faced by online FAQ systems, specifically in managing user inquiries. By utilizing BERT, we eliminate the need for users to have domain knowledge or for administrators to classify multiple inquiries. (2) We propose a higher accuracy inquiry classification model through contrastive learning that considers the similarity of questions and answers in loss function.



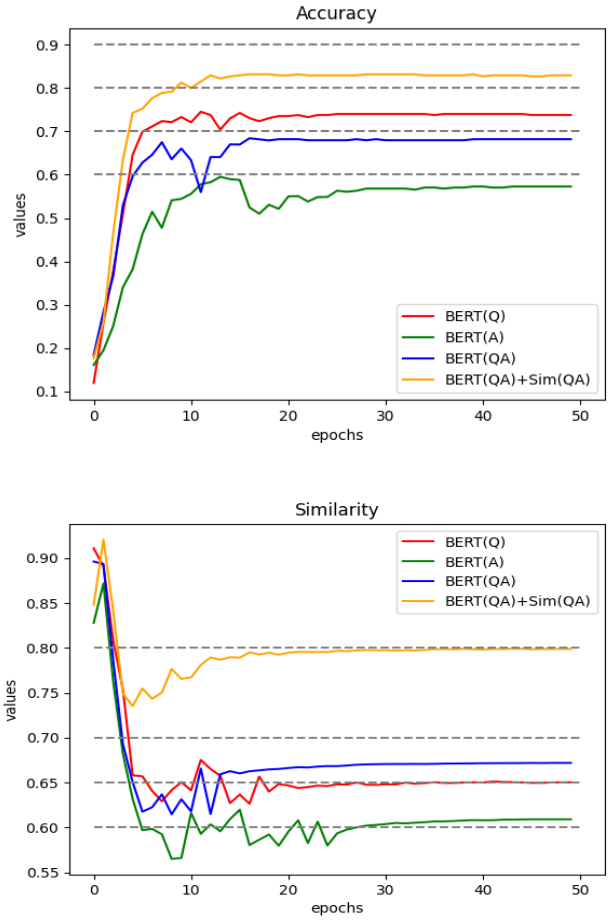
<Figure 2> Proposed model architecture. CLS(Q) refers to the CLS token obtained when the question is passed through the model.

2. Methodology

We collected 2576 question-and-answer pairs on cloud platforms, including our company. Each question is classified into one of 27 cloud-related categories. Since there is a limit to training deep learning with only this number of pairs, we used Korean comments BERT(KcBERT) pretrained language model to perform classification with fewer data [6].

Classifying the user's inquiry corresponds to a question among the question-and-answer pairs. However, we considered how we could use not only the questions but also the answers data we have. We thought that for more accurate inquiry classification, questions and answers should be represented as similar vectors within the same category.

Therefore, we have made the model perform two tasks in training. It uses a basic classification task and a task that makes the vectors of inquiries and answers similar. That is, as shown in <Figure 2>, the loss function for classification and the loss function using contrastive learning were



<Figure 3> Accuracy and similarity for the validation set of four experiments as training progresses.

combined and used as loss function for the model. Noted that in the case of contrastive learning, as mentioned in [5], the training was conducted with only positive pairs except for negative pairs because negative pairs can cause instability in model training and unnecessary computing time consumption.

3. Results

For the experiment setup, a learning rate of 0.0005, a warm-up ratio of 0.2, and a maximum length of the input token of 300 were used. In addition, the epoch is 50, the batch size is 32, and 20% of the total dataset is used as test sets. We conducted four kinds of experiments. The first is a task that predicts the categories of inquiries with only inquiries, the second is a task that conducts with only answers, and the third is a task that uses both inquiries and answers. Finally, the fourth is a task that uses both Inquiries and answers and uses contrastive learning to predict the category of inquiries. As shown in <Figure 3>, the fourth task, BERT(QA)+Sim(QA), showed higher accuracy than other methods while steadily increasing the similarity between questions and answers in validation set as training progressed. The results were the same in the test set. As shown in Table 1, BERT(QA)+Sim(QA) showed the highest accuracy as well as similarity.

<Table 1> Accuracy, similarity results of applying four tasks to the test set.

	Accuracy	Similarity
BERT(Q)	0.748	0.65
BERT(A)	0.518	0.605
BERT(QA)	0.61	0.63
BERT(QA)+Sim(QA)	**0.814**	**0.78**

4. Conclusion

This study proposed a model that uses BERT and contrastive learning to manage users' inquiries and automatically classify them into appropriate categories. The proposed approach eliminates the need for users to have domain knowledge and for administrators to classify multiple inquiries, improving the overall user experience of the FAQ system. The study evaluated the effectiveness of the proposed approach and compared it with other tasks. The results showed that the proposed approach had a greater accuracy with a 0.87 accuracy. The contribution of this study was to reduce the unnecessary time involved in classifying inquiries from the perspective of users and administrators facing the online FAQ system and the higher accuracy inquiry classification model through contrastive learning that considers the similarity of questions and answers in the loss function. In future research, it is recommended to investigate how the proposed model can be applied to other domains and languages and how it can be extended to handle more complex inquiries.

REFERENCES

[1] Devlin, Chang, Lee, & Toutanova. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding.*

[2] Shaw, Tsao, & Wu (2012). *A study of the application of ontology to an FAQ automatic classification system.*

[3] Damani, Narahari, Chatterjee, Gupta, & Agrawal. (2020). *Optimized transformer models for FAQ answering.*

[4] Mitra, Gupta, & Dandapat. (2020). *Transformer Models for Recommending Related Questions in Web Search*.

[5] Grill, Strub, Altché, … & Valko, (2020). *Bootstrap your own latent-a new approach to self-supervised learning.*

[6] Lee, J. (2020). *Kcbert: Korean comments bert.*