Preliminary Study on NLP technology for application of advanced MCRs

Taejin Kim [a*], Donghan Yoo [a], Geeyong Park [a], Kyungtae Lim [b], Seonggyu Lee [c]
[a]*Korea Atomic Energy Research Institute, 111, Daedeok-daero 989beon-gil, Yuseong-gu, Daejeon, Republic of Korea*
[b]*Hanbat National University, 125, Dongseo-daero, Yuseong-gu, Daejeon, Republic of Korea*
[c]*Snippod Inc., 17, Pangyoyeok-ro 192beon-gil, Bundang-gu, Seongnam-si, Gyeonggi-do, Republic of Korea*
[*]*Corresponding author: taejinkim@kaeri.re.kr*

1. Introduction

As advanced digital technologies are applied to main control rooms (MCRs) in nuclear power plants (NPPs), these are helpful to reduce workload and to relieve stress for MCR operators to some degree. However, it does not bring a huge change for the number of operators in MCRs.

In advanced MCRs, putting natural language processing (NLP) technology to human-system interface (HIS) design can be a useful tool to minimize the number of operators by replacing the traditional tasks of operators with NLP technology. Thus, we review the latest NLP technology and introduce the NLP-based system adequate for advanced MCRs.

2. Natural Language Processing(NLP) Technology

In this section, the state-of-the-art NLP technology which is BERT is introduced. The transfer learning, which is for enhancing performance of a model with pre-trained model, is described.

2.1 State-of-the-art NLP technology: BERT

As it has been showing a good performance to utilize large-scale pre-trained models for NLP, it is preferred to use pre-trained models in AI field [1]. Bidirectional Encoder Representations from Transformers (BERT) is one of the most state-of-the-art pre-trained models in the last few years.

BERT was introduced by Jacob Devlin and his colleagues from Google [2]. It is pre-trained with two unsupervised steps: masked language modelling (MLM) and next sentence prediction (NSP). For a bidirectional training, 15% of input tokens are masked and then the model is trained to predict the original tokens during the MLM step. In order to make the model understand sentence relationship, a next sentence prediction task is trained in the NSP step.

BERT is considered a game changer in NLP field due to its benefits and benefits of BERT is as follows [2]:

- Able to train large amount of data with unsupervised learning
- High performance with a deep bidirectional architecture over unidirectional models
- Easy to acquire desired models by fine-tuning pre-trained models with task-specific training data

2.2 Transfer Learning

Jacob Devlin and his colleagues [2] also introduce the fine-tuning BERT. In fine-tuning step, training inputs and outputs are selected dependent on its purpose, and then the pre-trained BERT are trained with them end-to-end.

That is, transfer learning is to add specific input and output data related to the purpose of the model into the pre-trained model. In this way, leveraging prior knowledge can be utilized to enhance performance of the model. The concept of the transfer learning is simply represented in Fig. 1.

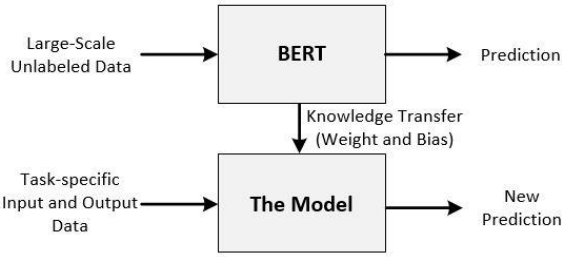

Fig. 1. Transfer Learning: The relationship between BERT and the transfer learning model

A taxonomy for transfer learning is suggested by Ruder [3]. It is broadly classified with inductive and transductive transfer learning. Inductive transfer learning is defined as having different source and target task, and it is divided into multi-task learning and sequential transfer learning whether the tasks are learned simultaneously or sequentially. Transductive transfer learning is in cases of having same source and target, and it is also divided into domain adaptation and cross-lingual learning whether the datasets come from different domains or different languages.

3. NLP-based System for advanced MCRs

In this section, the role and purpose of the NLP-based system are described. Also, benefits of the NLP-based system are suggested when the NLP-based system is utilized during communication of MCR operating team.

3.1 Description of the NLP-based system

For South Korea, the configuration of operating team is comprised of a shift supervisor (SS), a reactor operator (RO), a turbine operator (TO), an electric operator (EO), and a shift technical advisor (STA). For

reference, the term of board operators (BOs) is referred to RO, TO, and EO. A NLP-based system is designed to replace most of BO's tasks in advanced MCRs. In advanced MCRs, the NLP-based system processes and analyzes voice information of a SS and conveys inquiries or commands of the SS to the existing information processing system. Also, the NLP-based system extracts the information from the existing information processing system for the response to the SS's inquiries and for the action report to the SS's commands, and provides the response or the action report to the SS. It is comprised of four elements in Fig. 2.:

1) Voice Receiver
2) Voice Sender
3) Speech Recognition AI Module
4) Natural Language Processing AI Module

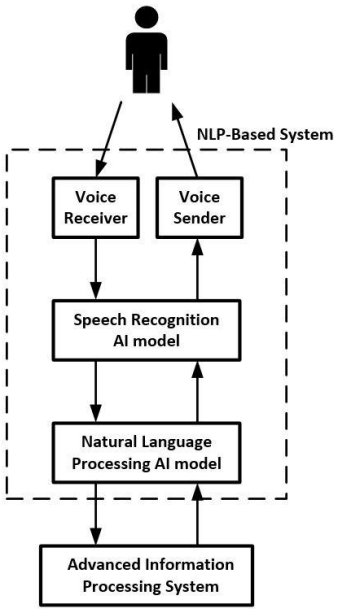

Fig. 2. Configuration of the NLP-based system

When it comes to tasks of operators in advanced MCRs, Kim and his colleges classify them into 3 factors: cognitive tasks, communitive tasks, and operational tasks. The definition of each activity is as follows and classification of each activity [4]:

1) Cognitive Task: The cognitive activities of operators conducted for operation and situation handling of a power plant
2) Communitive Task: The communicative activities of operators conducted for operation and situation handling of a power plant
3) Operational Task: The behavior that operators take to conduct the cognitive activities

Looking at each activity, a SS mostly takes care of cognitive activities and BOs conduct COMPARE, DIAGNOSE, and MONITOR among cognitive activities [4]. The NLP-based system can compare two or more entities upon a SS's request, and monitor system status, but a SS should diagnose the cause of a condition based on information provided by the NLP-based system if there are no BOs in an operating team. For communitive activities, BOs mainly communicate with a SS and partly interact with on-site operators. The NLP-based system can substitute BOs' roles during the conversation with a SS, but a SS should take BOs' roles to communicate with on-site operators. Operational tasks are divided into NAVIGATION, MANIPULATION, and CONFIRMATION [4]. Among them, BOs mainly conduct NAVIGATION and MANIPULATION and these activities can be conducted by the NLP-based system under the supervision of a SS.

In order to replace certain tasks of BOs with the NLP-based system, the design of Natural Language Processing AI module in the NLP-based system is key to success. The pre-trained BERT will be utilized to design the Natural Language Processing AI module, and the pre-trained BERT will be fine-tuned for purpose. Based on the output of the fine-tuning model, the answer will be generated in a rule-based way which is determined under the consideration of tasks and communication way of operators. The architecture of the Natural Language Processing AI module is represented in Fig. 3.

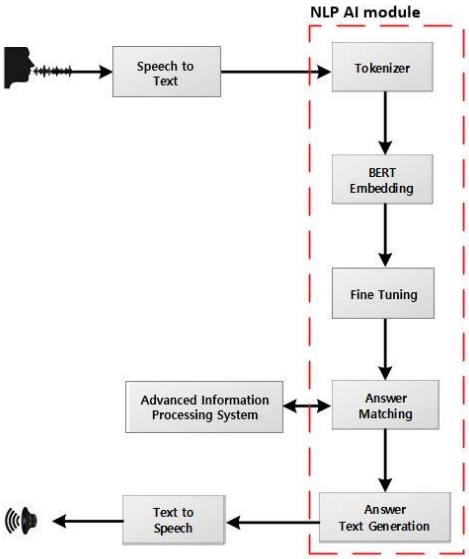

Fig. 3. Architecture of the NLP AI module

3.2 Advantages of the NLP-based system

Main purpose for application of the NLP-based system is to minimize the number of staff in advanced MCRs. It will improve economic efficiency and competitiveness of nuclear power plants. Also, small modular reactors (SMRs) are actively researched these days, the need gradually increases to find an optimal way to operate reactors. There is a consensus that the

current number of operators in MCRs is not acceptable for SMRs. Thus, application of the NLP-based system into MCRs can provide one of measures to solve this problem.

Since most of MCRs are digitalized, the NLP-based system can be effectively compatible with the existing system. Also, simple and fast communication can be still maintained because it is still possible to communicate verbally between a SS and the NLP-based system.

In addition, the NLP-based system can contribute to prevent errors occurred by board operators. As MCRs have been digitalized, board operators are required to do secondary tasks, and it provokes errors, such as operation omission, inadequate operation, and delayed operation [5]. Also, there exists human errors caused by human intrinsic characteristics of board operators, such as fatigue, embarrassment, etc., under the condition of excessive stress and workload. The NLP-based system can execute SS's requests without errors, provided that a SS's requests are correctly understood, and is not influenced by stressful or overburden circumstances.

4. Conclusions

This paper conducts a preliminary study to replace some tasks of board operators with the natural language processing (NLP) technology. The state-of-the-art NLP technology which is BERT is introduced and the pre-trained BERT will be utilized to design the NLP-based system. The tasks of board operators which can be replaced by the NLP-based system are analyzed. The analysis will be used to design a method for transfer learning of the pre-trained BERT model in the future. Application of the NLP-based system into advanced MCRs will contribute to reduce staffs and human errors.

REFERENCES

[1] Xu Han, Zhengyan Zhang, et al., Pre-trained models: Past, present, and future, AI Open 2, pp.225-250, 2021.
[2] Jacob Devlin, Ming-Wei Chang, et al., BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, 2019.
[3] Ruder, S., Neural Transfer Learning for Natural Language Processing, Ph.D. thesis, National University of Ireland, Galway, 2019.
[4] Kim, S., Kim, Y., Jung, W., Operator's cognitive, communicative and operative activities based workload measurement of advanced main control room. Annals of Nuclear Energy 72, p. 120–129, 2014.
[5] Lee, S., Kim, J., JANG, S., Human Error Mode Identification for NPP Main Control Room Operations Using Soft Control. Journal of Nuclear science and technology, Vol. 48, No. 6, p. 902-910, 2011.