[Transactions of the Korean Nuclear Society Spring Meeting Jeju, Korea]
May, 18-19, 2023

Exploring Methodologies to Increase the Reliability of Data for Verifying Nuclear Power Plant Artificial Intelligence Software

Da-Young Lee*, Han-Gil Lee, Jeong-Hun Lee
FNC TECHNOLOGY CO.,LTD/ 13 Heungdeok 1-ro, 32F, Giheung-gu, Yongin-si, Gyeonggi-do, 16954, Korea (*ekdud93@fnctech.com)

Introduction

- The impact of artificial intelligence technology on the industry is gradually increasing.
- AI should have sufficient reliability in the nuclear field.
- In this paper, we conducted a study on software verification from the perspective of data balance to apply artificial intelligence-based software to MMIS

Background

Explainable artificial intelligence methodology

- Fig1. shows the process of AI detecting defects in self-produced nuclear reactor components, using the Yolov7 model with CNN algorithm applied for data training.

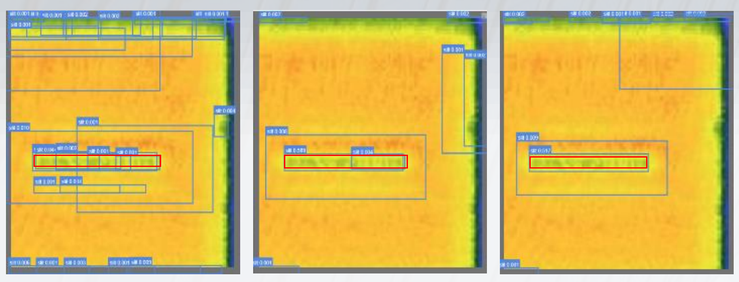

Fig 1. Detection Data-1

- Fig 2. and Fig 3. are cases of verifying the reliability of AI through models designed with a focus on XAI.

Fig 2. Detection Data-2 / Fig 3. Detection Data-3

- By visualizing the black box area in this way, it is possible to enhance the explanatory power for quality verification of the AI module.

Appropriate data balance methodology

- Pursue the reliability of the data used for AI learning itself.
- Fig 4. shows the performance of AI results that detected "Load Trucks" improved when trained on 2,000 data with a proper data balance compared to results trained on 50,000 normal data.

Fig 4. Detection Data-4

Designing AI data balance for NPP

- Our goal is to apply the software design methodology proposed by IEEE 1012 to the necessary data balance design.

Table 1. IEEE1012 Software Design Phase Output

SDLC(System Development Life Cycle)	Deliverables
Software Design	SDD, CT/IT Plan, CT/IT/ST Procedure, RTM, V&V Report

- Table2. shows the parameters defined for the analysis of AI reliability in TTA(Telecommunications Technology Association design procedure)

Table 2. Data Quality Metrics

No	Testing Items	means
1	Evaluation Data Meta Attribute Count, (MC)	Meaning property complexity of data
2	Evaluation Data Meta Attribute Value Count, (MVC)	Indicates the complexity of the data attribute value
3	Balanced Evaluation Dataset Count,(EDC)	Minimum number of assessments required to balance data
4	Evaluation Applied Dataset Count, (AC)	Number of datasets used to measure the accuracy of actual AI-based software
5	Coverage of Evaluation Applied Dataset, (CAD)	This refers to the ratio of the actual data set to the total data set required for the evaluation

A: The reliability of artificial intelligence software using a balance-based dataset

$$B = \text{mAP(Mean Average Precsion)} = \frac{1}{n}\sum_{i=1}^{N} AP_i \qquad (1)$$

$$A = B*(AC/EDC)*100 \qquad (2)$$

- As the ratio of the number of datasets used for evaluation of the minimum evaluation dataset(EDC) to the number of datasets applied for evaluation(AC), increases in equation(2) the reliability is expected to improve. As a result, come to a conclusion to test case problem coverage.

Conclusion

- For software reliability verification, the black box characteristics of AI and the data quality aspect should be considered.
- The design criteria of AI for reliable data balance, which meets the regulatory requirements of the MMIS is considered.
- The methodology to improve the reliability of AI through the process of expanding test data coverage will be studied.

Acknowledgements

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2022-00144521).