*Transactions of the Korean Nuclear Society Autumn Meeting*
*Changwon, Korea, October 20-21, 2022*

# Application of Supervised Learning Algorithms for Classification of Level of Partial Defect in Spent Nuclear Fuel with Scintillator Based Partial Defect Detector

Yeongjun Kim[a], Man-Sung Yim[a*]
[a]*Department of Nuclear and Quantum Engineering, KAIST 291 Daehak-ro, Yuseong-gu, Daejeon, 34141, Republic of Korea*
[*]*Corresponding author: msyim@kaist.ac.kr*

## 1. Introduction

The International Atomic Energy Agency (IAEA) requires various types of safeguards to prevent nuclear proliferation and the diversion of nuclear materials [1]. Nuclear safeguards consist of containment, surveillance, accountancy and the inspection of nuclear materials [2]. Through the accountancy and the inspection, IAEA can identify whether nuclear material has been diverted and whether the real quantities of nuclear material correspond with the reported data. Typically, the quantity verification is performed in the spent fuel pool before the spent fuel is permanently disposed or before they become difficult-to-access. The safeguard criterion is when half the fuel in a fuel assembly is unaccounted for. Various types of NDA detectors are available to satisfy this criterion but the function of most of them is limited to identify the presence of the diversion [3]. Machine learning techniques can help these NDA detectors to know how much and where the fuel diversion occurred. Previous research at SCK CEN Belgian Nuclear Research Center examined typical machine learning algorithms such as the decision-tree or k-nearest neighbors with the Fork detector(FDET), Self-Indication neutron resonance densitometry(SINRD), and Partial defect tester(PDET) to classify the level of fuel diversion [4]. It was shown that the machine learning algorithms can be applied to improve the safeguards capabilities.

To overcome the limitations of existing NDA techniques, a scintillator based partial defect detector(SPDD) is under development in KAIST [5]. It measures the gamma ray emitted from the fission products in the spent fuel. It decides the presence of partial defect with identifying decreases of signal based on the estimated signal from the normal spent fuel with uniformly distributed sensor pixels over the fuel assembly. To be competitive with current NDA techniques used for safeguards, this detector needs to be more versatile, precise and autonomous detector that does not even require any engagement of the inspectors. Hence, the applicability of the machine learning algorithms was assessed with the SPDD. The optimal classification model was found among five supervised learning algorithms. The prediction accuracy was evaluated based on the measurement with randomly distributed partial defect scenarios.

## 2. Methods

The SPDD measures gamma rays from the spent fuel and it was designed to be inserted into the guide tubes of the fuel assembly which is very similar to the concept of the PDET [6]. However, it has fundamental limitation that the detection capability of the SPDD is highly dependent on the types of fuel assemblies because they have different distribution of the guide tubes. Thus, the location of the sensor pixel is changed from the center of the guide tube to above the fuel assembly. Figure 1 shows the optimized distribution of the sensor pixels in the Westinghouse 17x17 fuel assembly.

To apply the classification algorithms, large amount of measurement data needs to be obtained for the training and test of the model. MCNPX was used for the simulation of the spent fuel inspection by the SPDD [7]. The gamma source such as the intensity and the energy spectrum of the spent nuclear fuel was generated with using the ORIGAMI module of the SCALE code [8]. The target fuel assembly type was assumed as standard Westinghouse 17x17 because the design of SPDD was optimized with it. The operation histories such as the initial enrichment, discharge burnup, and the cooling time in the pool were assumed as the averaged value of the real 17x17 spent fuels stored in South Korea as shown in table 1. The fuel diversion scenarios for the training and test data were based on the randomly distributed dummy rods. A dummy rod is assumed to be a fresh fuel or a rod filled with SS304.

Five supervised classification models were used. For the classification of the ratio of the defects, multinomial Logistic Regression, Support Vector Machine, k-Nearest Neighbor, Naïve Bayes, and Decision Tree models were utilized. The Scikit-learn library from the Python framework was used to apply the models to obtained data [9]. The input parameters are the estimated gamma count rates from the sensor pixels. There are 24 and 25 input features which can be obtained from the measurement for the original and optimized design. The target class to classify is the ratio of number of defects to the entire number of fuel rods in a single fuel assembly. The interval of each classes, number of data and the label is shown in table 2.

The accuracy scores of each model was compared to find the optimal model with varying the forms of the features and classes.

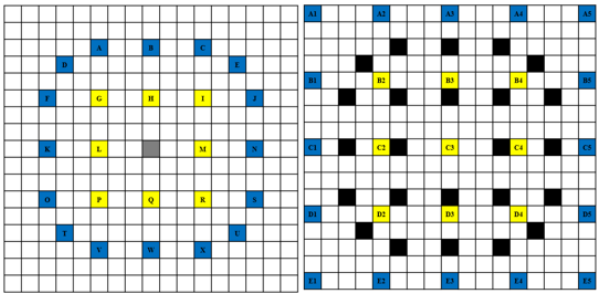

Fig. 1. Location of the sensor pixels of SPDD in Westinghouse 17x17 type (left ; original, right ; optimized, blue ; peripheral, yellow ; central)

Table. 1. Assumption of operation histories of spent fuel for gamma source

| Variable | Range |
|---|---|
| Initial enrichment | 4.5wt% of U-235 |
| Discharge burnup | 40GWD/MTU |
| Cooling time | 1 years |

Table. 2. Target classes for classification

| | Class 1 | Class 2 |
|---|---|---|
| Interval | 2%(0~8%) | 4%(0~8%) |
| Label | 1, 2, 3, 4, 5 | 1, 2, 3 |
| Number of data for each class | 100 | 100 |

## 3. Results

To find the optimal model, the accuracy score metric was used and compared for each model. The k-fold cross validation method was also used for statistical analysis of the accuracy. It is useful to improve the credibility of the accuracy when the number of the data is limited. The ratio of test data to total dataset was fixed to 20%. For the classification of the 2% defect interval with the entire sensor pixels, except the Decision Tree model, other four models showed relatively good prediction performance as shown in figure 3. Because of the large number of the input features, the nodes and branches of tree became complicated. Hence, the tree model was inappropriate for the data obtained by the SPDD. The accuracies of other four model was significantly improved by the design optimization because the sensor pixels were more uniformly deployed over the fuel assembly. When the interval of each class become 4% of defect, the overall accuracies of five models were increased as shown in figure 4. It was shown that the 4% of defect can be easily classified even with the original design of SPDD. With the data obtained by the optimized design, the accuracy reached 1 with the multinomial LR, SVM and the Naive Bayes models. The accuracy scores were also analyzed with varying the number of input features. When only the peripheral sensors were used, the overall accuracies were slightly decreased compared to those obtained with the entire sensor pixels. However, the four models except the Decision Tree showed still quite good performance with the optimized design. When only nine central sensor pixels were used, the accuracies were significantly decreased compared to those obtained with the entire or peripheral sensor pixels. Hence, the peripheral sensor pixels are relatively more important than those located in the central region.

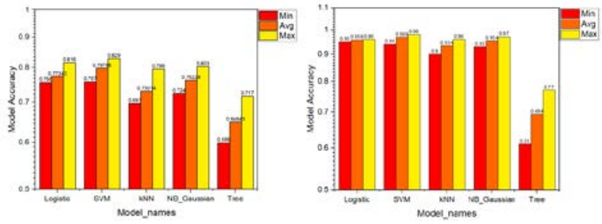

Fig. 3. Accuracy scores for each model with Class 1 and entire sensor pixels (left ; original, right ; optimized)

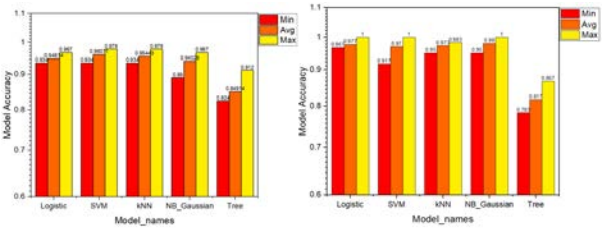

Fig. 4. Accuracy scores for each model with Class 2 and entire sensor pixels (left ; original, right ; optimized)

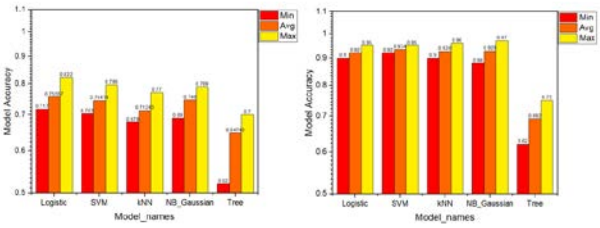

Fig. 5. Accuracy scores for each model with Class 1 and peripheral sensor pixels (left ; original, right ; optimized)

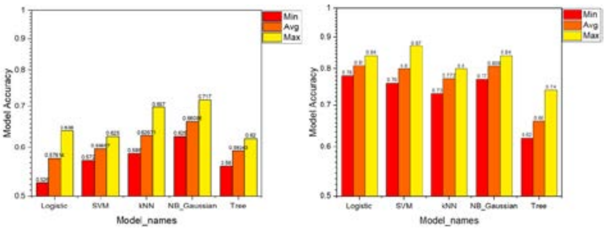

Fig. 6. Accuracy scores for each model with Class 1 and central sensor pixels (left ; original, right ; optimized)

## 4. Conclusions

To develop an autonomous partial defect detector, typical supervised learning algorithms were applied to predict the presence and the degree of partial defects. It was shown that 2% level of defect can be classified with using the classification models. With the optimized design of SPDD, four models(multinomial LR, SVM, kNN, and gaussian NB) show more than 95% of prediction accuracies. To predict not even the presence and the degree of the partial defect but also the distribution and exact location of the defect, deep learning algorithms such as CNN or GAN will be applied in further research.

## ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2016R1A5A1013919).

## REFERENCES

[1] Chen, J. D., Gerwing, A. F., & Lewis, P. D. (2001). Long-cooled spent fuel verification using a digital Cerenkov viewing device.

[2] IAEA, International Safeguards in the Design of Nuclear Reactors, IAEA Nuclear Energy Series NP-T-2.9, IAEA, 2014.

[3] International Atomic Energy Agency (IAEA), Safeguards techniques and equipment: 2011 edition, IAEA-NVS-3, 2011.

[4] Rossa, R., Borella, A., & Giani, N. (2020). Comparison of machine learning models for the detection of partial defects in spent nuclear fuel. Annals of Nuclear Energy, 147, 107680.

[5] Lee, H., & Yim, M. S. (2020). Investigation of a fast partial defect detection method for safeguarding PWR spent fuel assemblies. Annals of Nuclear Energy, 144, 107496.

[6] Ham, Y.S., Sitaraman, S., 2011. Partial Defect Tester: A novel approach to detect partial defects in pressurized water reactor spent fuel. Nuclear Technol. 175, 401–418.

[7] Pelowitz, D.B.,2011. MCNPX User's Manual Version 2.7.0, LA-CP-11-00438, pp. 107– 113.

[8] Rearden, B. T., & Jessee, M. A. (2018). SCALE code system (No. ORNL/TM--2005/39-V-6.2.3). Oak Ridge National Laboratory (ORNL).

[9] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.