Transactions of the Korean Nuclear Society Virtual Spring Meeting
July 9-10, 2020

Feasibility study on machine learning algorithm in nuclear reactor core diagnosis

Hanjoo Kim [a], Dongmin Yun [a], Ho Cheol Shin [b], Sang Rae Moon [b], Deokjung Lee [a*]
[a] *Department of Nuclear Engineering, Ulsan National Institute of Science and Technology UNIST-gil 50, Ulsan, 44919*
[b]*Core and Fuel Analysis Group, Korea Hydro & Nuclear Power Central Research Institute (KHNP-CRI), Daejeon, 34101*
[*]*Corresponding author: deokjung@unist.ac.kr*

1. Introduction

Due to current trend of light water reactors (LWRs), high thermal power, obsolescence, long-cycle, and water chemistry management, the importance of nuclear reactor core monitoring has been increased since early diagnostics of core abnormal state, and follow-up measures on it can reduce costs caused by the abnormal state. In this study, a reactor core diagnostics program based on machine learning (ML) algorithm is under development to improve the current reactor core monitoring system which is based on operator's proficiency. Due to the limitation of obtaining enough operation data to train and test a ML algorithm classifying various reactor core conditions, a nuclear core analysis code, RAST-K, is employed to generate dataset. Reactor abnormal conditions such as CRUD induced power shift (CIPS) and control rod mis-location are modeled, and each model with randomly perturbated input parameters are labeled. Supervised learning is performed, such that control rod positions and detector signal data with corresponding label of reactor core conditions is used as training data of the ML models. Since the final goal of the ML model is being implemented in reactor core monitoring system, detector signal and control rods data which also can be observed by reactor operators are used as input of the ML model.

2. Training data generation

2.1. Data generation system

Training an ML model requires huge amount of dataset to achieve high performance of it. For this reason, training dataset generation system named AUTOGEN has been established. The reactor analysis code RAST-K [1] is embedded in the system to compute detector signals [2] used in training of ML models. The data generation system is written by Python language, and it works with three steps as:

1) Generation of abnormal core model and RAST-K input by randomly perturbing input parameters
2) Running RAST-K for all generated input files
3) Extraction of target output parameters from output files and write it in single train dataset file with "csv" format.

The structure of data generation is shown in Figure 1. In this study, 4th cycle of OPR1000 type reactor with full core model is used as base model of the dataset generation.

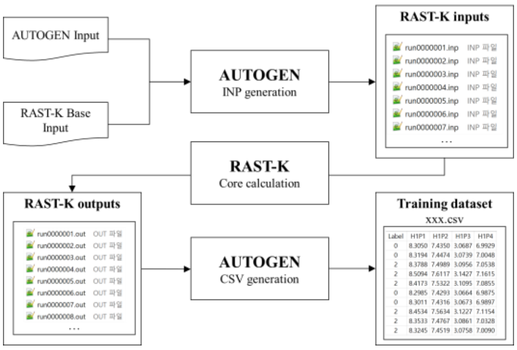

Figure 1. Data generation system structure

2.2. CRUD occurrence core model

To model the core model with CRUD accumulation, a simple CRUD accumulation model is implemented in the RAST-K. The CRUD thickness of a fuel assembly node is defined with CRUD multiplication factor (α) and subcooled nucleate boiling mass (SNB_{mass}) as follow:

$$\tau_{CRUD} = \alpha \times SNB_{mass} \ . \tag{1}$$

The boron number increase at i^{th} depletion burnup step is

$$dN_B = ND_B \times dV_{CRUD} \ (\text{if } \tau_{CRUD} > \tau_{threshold}) \ , \tag{2}$$

where the CRUD volume increase is determined with the node height (h), the CRUD porosity ($k_{porosity}$), and the outer radius of the CRUD layer (r_{CRUD}) as

$$dV_{CRUD} = 2\pi h(1-k_{porosity})(r_{CRUD,i}{}^2 - r_{CRUD,i-1}{}^2) \ , \tag{3}$$

the boron number density is defined with reference boron number density (ND_{ref}) at 800 ppm, and critical boron concentration (CBC) as follow

$$ND_B = ND_{ref} \times (CBC/800) \ . \tag{4}$$

Since it was analyzed that the minimum thickness of the CRUD for CIPS is 30 μm [3], threshold thickness ($\tau_{threshold}$) for boron holdup is 30 μm.

The core average CRUD thickness is shown in Figure 2 and it shows growth of CRUD during the operation.

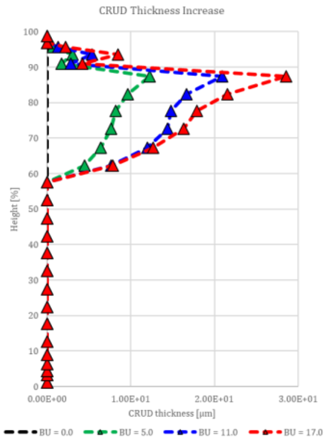

Figure 2. CRUD thickness increases

Figure 3 shows comparison of ASI between core models with CRUD and without CRUD. Since boron holdup requires minimum thickness of CRUD, ASI starts to change from 2.0 GWd/MTU. Since CRUD accumulation and boron holdup appear at upper region of fuel assemblies (FAs), ASI is higher with CRUD, and it is lower at the end of cycle (EOC) due to low burnup of fuel at upper region.

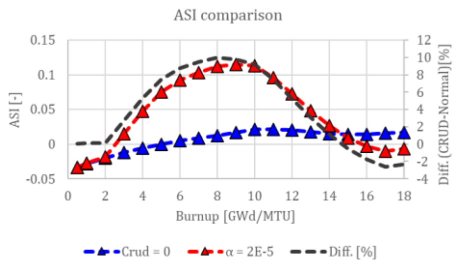

Figure 3. ASI comparison with CRUD occurrence

In generation of core models with CRUD occurrence, CRUD multiplication factor (α) is perturbed to generate various core conditions caused by CRUD accumulation. 1% difference of ASI is selected as a criterion for labelling CIPS data. Once CIPS occurs at certain burnup step, core operation data from the burnup step are labelled as CIPS. The procedure to generate the CRUD accumulated model is as follow:

1) Randomly select CRUD multiplication factor with uniform distribution from 1E-6 to 1.5E-5.
2) Perform steady state depletion calculation from 0 to 17.0 GWd/MTU by using RAST-K.
3) Labelling the data as CIPS or normal if ASI difference exceeds 1%.
4) Repeat 1) ~ 3).

In this study, 10,000 input files with 21 burnup steps were generated and calculated, giving 210,000 snap-shop data. In the training and testing a ML model which classifies CRUD occurrence, 30,000 data was used for middle of cycle (MOC) and end of cycle (EOC), respectively. The core burnup for MOC data is 5.0 ~ 7.0 GWd/MTU, and the core burnup for EOC data is 15.0~17.0 GWd/MTU.

Table 1. Dataset for CRUD occurrence model

Burnup	Power	Total	Normal	Abnormal
MOC	100%	30,000	18,844	11,156
EOC	100%	30,000	12,969	17,031

2.3. Control rods mis-location core model

Reactor core abnormal condition with control rods mis-location is modeled. The model represents a core condition where a CEA position is different with other CEAs in the same sub-group, and thus, it can cause power tilt during the operation. The core model is labeled as control rod mis-location if the difference of CEA position exceeds 8.52 cm. The criterion of 8.52 cm is defined with minimum rod worth of a regulating bank ($\rho_{R,\min}$), acceptable design uncertainty of control rod worth (10% or 100 pcm), maximum differential rod worth of a shutdown bank ($\frac{d\rho}{dh}_{S,\max}$) as

$$h_{criteria} = 10\% \times \rho_{R,\min} \Big/ \frac{d\rho}{dh}_{S,\max} \quad . \qquad (5)$$

Since 25.15 cm of the difference is a criterion used in OPR1000 reactor to reactor shutdown, 8.52 cm of the difference in this study is adjustable to develop a ML model for early diagnostics. Control rods in OPR1000 reactor are driven by a motor having steps of 1.905 cm [4] and thus, control rod position is determined by sampling number of control rod steps. The procedure to generate control rod mis-location model is:

1) Randomly sample the regulating bank position satisfying PDIL to determine normal core condition.
2) Randomly select a CEA out of 73 CEAs to perturb control rod position.
3) Sample the number of steps to insert or withdraw
4) If the sampled step is more than 5 steps, the core model is labeled as control rod mis-location.
5) Perform core calculation by using RAST-K.
6) Repeat 1) ~ 5).

10,000 data of control rod mis-location model were generated for BOC, 50,000 data for MOC and EOC were generated. The core operating conditions are hot full power (HFP), 80%, and 60% core power.

Table 2. Dataset for control rod mis-location model

Burnup	Power	Total	Normal	Abnormal
BOC	100%	10,000	5,528	4,472
	80%	50,000	27,288	22,712
	60%	50,000	27,061	22,939
MOC	100%	10,000	5,587	4,413
	80%	50,000	27,123	22,877
	60%	50,000	27,156	22,844
EOC	100%	10,000	5,552	4,448
	80%	50,000	27,204	22,796
	60%	50,000	27,093	22,907

3. Training ML models

3.1. Data description

Dataset with “csv” format is used for training the ML model. Since ML model with supervised learning is target model of this study, label is included in the dataset. The first column of the dataset indicates the core states (label) and the second to end column represents detector signal data in the model. Figure 4 shows example of training data file generated by the AUTOGEN. Each row represents information of a core model and first column of the data is label for the core model, and rest are control rod (CR) position and in-core instrument (ICI) signal. The control rod position in X data is nominal position.

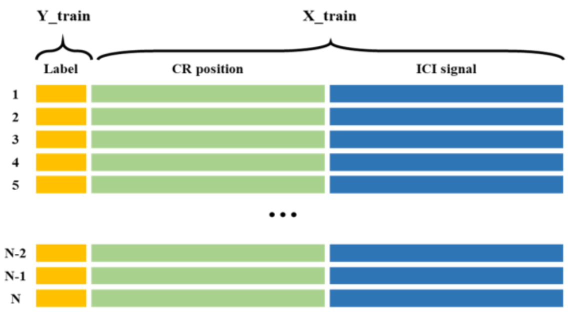

Figure 4. Example of training data file

Full core of OPR1000 type reactor is base core model of this study and it includes 73 CEAs and 225 detector signals (Radially 45 * axially 5). Figure 5 and 6 show position of CEAs and ICI FAs. Each ICI has 5 in-core detectors located at 10%, 30%, 50%, 70% and 90% of height, respectively.

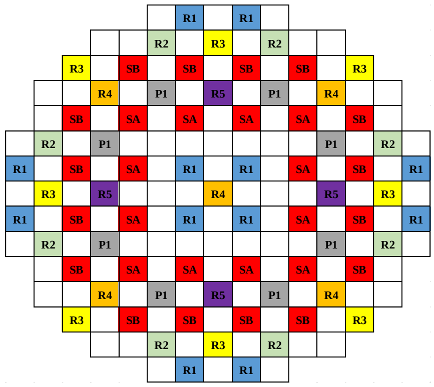

Figure 5. CEAs for OPR1000 type reactor

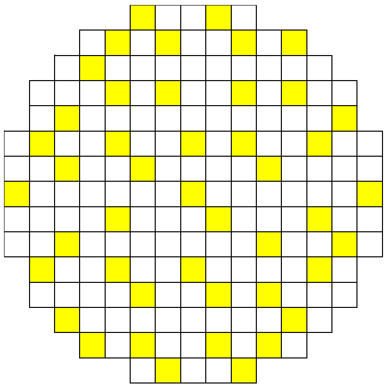
Figure 6. ICI assemblies for OPR1000 type reactor

3.2. ML model description

Naïve Bayes (NB), Support Vector Machine (SVM), Deep Neural Network (DNN) and Random Forest (Random Forest) model were built and each model represents a stochastic model, a linear classification model, a neural network model and an ensemble model. Data was normalized such that it has 0 mean and unit variance. The normalized input data with dimension of 293 is used as input of the ML models. The generated datasets of 390,000 were used to train, validate and test the ML models, and 60% of the dataset is used as trainset, and validating and testing used 20% of the dataset.

Table 3 shows summary of the ML model to build.

Table 3. Summary of ML model

Model	Description
NB	Stochastic model: Gaussian
SVM	Kernel: Linear Max. iteration: 3,000 C: 1.0
DNN	Number of hidden layers: 4 (D = 64) Activation function: ReLu Number of trainable parameters: 31,746 Optimizer: Adam Epochs: 300
RF	Number of trees: 200 Depth of tree: Max.

4. Results

4.1. Classification for reactor core state

Prediction accuracy of the ML models were compared for CRUD occurrence and CR mis-location. Training on binary classification (normal vs. abnormal) was performed and its prediction accuracy was tested. The accuracy of the model is defined with true positive (*TP*), true negative (*TN*), false positive (*FP*) and false negative (*FN*) as follow:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad . \qquad (6)$$

Binary classification prediction results for the ML models are shown in Table 4 and Table 5. All ML models well predicts CRUD occurrence showing 90% prediction accuracy.

Table 4. Prediction accuracy [%] for CRUD occurrence

	NB	SVM	DNN	RF
MOC	92.90	99.95	99.98	99.86
EOC	97.43	99.96	99.96	99.98

The prediction accuracy for CR mis-location reduces as core power decreases. As power decreases, more regulating bank positions are possible since power dependent power insertion limit (PDIL) is loosen at lower power. Since that, as power decreases, the ratio of unseen data in test data increases, decreasing the accuracy.

Table 5. Prediction accuracy [%] for CR mis-location

	Power	NB	SVM	DNN	RF
BOC	100%	79.50	87.00	100.00	100.00
	80%	52.74	49.01	95.52	100.00
	60%	54.43	53.08	89.50	88.35
MOC	100%	79.40	83.65	100.00	100.00
	80%	51.97	47.09	96.67	92.87
	60%	55.50	51.62	91.11	89.12
EOC	100%	78.40	39.55	100.00	100.00
	80%	56.19	47.30	98.37	93.65
	60%	55.54	47.68	93.69	89.37

The synthesis operation data generated by the RAST-K does not include noise of the signal. To test feasibility of the ML models in real NPP data, 1% of artificial noise was applied to ICI signal. The noise has gaussian distribution with 1% standard deviation of the signal. The prediction accuracy of the ML models for CR mis-location with 1% ICI signal noise is shown in Table 6. By adding noise in ICI signal, prediction accuracy reduced comparing to results in Table 7. Because ICI signal is the only parameter representing power distribution in training data, noise of it reduces the prediction accuracy.

Table 6. Prediction accuracy [%] for CR mis-location with 1% of ICI signal noise

	Power	NB	SVM	DNN	RF
BOC	100%	79.50	87.00	100.00	100.00
	80%	52.74	49.01	95.52	100.00
	60%	54.43	53.08	89.50	88.35
MOC	100%	79.40	83.65	100.00	100.00
	80%	51.97	47.09	96.67	92.87
	60%	55.50	51.62	91.11	89.12
EOC	100%	78.40	39.55	100.00	100.00
	80%	56.19	47.30	98.37	93.65
	60%	55.54	47.68	93.69	89.37

Comparing the prediction results of tested ML models, DNN and RF model shows better accuracy comparing to NB and SVM.

5. Conclusion

In this study, framework for developing ML model in reactor core diagnostics has been established with three steps: 1) Establishment of dataset generation system based on RAST-K, 2) Generation of data for CRUD occurrence and CR mis-location core models and 3) Building and training of ML models. In this framework, the generated dataset was applied to train various ML models such as Naïve Bayes(NB), Support Vector Machine(SVM), Deep neural Network(DNN) and Random Forest(RF). Comparison of the results showed that DNN and RF model has better performance in reactor core diagnostics than other ML models. With high prediction accuracy of DNN and RF model, one can conclude that the application of a ML model in reactor core diagnosis is feasible. In the future, model data will be generated to reduce ratio of unseen data, additional parameters will be used as trainset for better performance. More reactor core models with abnormal condition will be generated for real NPP application.

Acknowledgement

This work was supported by KOREA HYDRO & NUCLEAR POWER CO., LTD (No. 2018-Tech-07)

REFERENCES

[1] Choe, Jiwon, et al. "Verification and validation of STREAM/RAST-K for PWR analysis." Nuclear Engineering and Technology 51.2:356-368,2019.
[2] Khoshahval, Farrokh, et al. "Vanadium, rhodium, silver and cobalt self-powered neutron detector calculations by RAST-K v2. 0." Annals of Nuclear Energy 111:644-659, 2018.
[3] Henshaw, Jim, et al., A model of chemistry and thermal hydraulics in PWR fuel CRUD deposits., Journal of nuclear materials 353.1-2:1-11, 2006.
[4] Yu, Je-Yong, et al., Sensory systems for a control rod position using reed switches for the integral reactor, 2007.