Nuclear Engineering and Technology 54 (2022) 3620–3630

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original Article

Consistency check algorithm for validation and re-diagnosis to improve the accuracy of abnormality diagnosis in nuclear power plants

Geunhee Kim, Jae Min Kim, Ji Hyeon Shin, Seung Jun Lee*

Ulsan National Institute of Science and Technology, 50, UNIST-gil, Ulsan, 44919, Republic of Korea

ARTICLE INFO

Article history:
Received 7 February 2022
Received in revised form
3 May 2022
Accepted 30 May 2022
Available online 2 June 2022

Keywords:
Nuclear power plant
Abnormality diagnosis
Diagnostic validation
Consistency check
Re-diagnosis

ABSTRACT

The diagnosis of abnormalities in a nuclear power plant is essential to maintain power plant safety. When an abnormal event occurs, the operator diagnoses the event and selects the appropriate abnormal operating procedures and sub-procedures to implement the necessary measures. To support this, abnormality diagnosis systems using data-driven methods such as artificial neural networks and convolutional neural networks have been developed. However, data-driven models cannot always guarantee an accurate diagnosis because they cannot simulate all possible abnormal events. Therefore, abnormality diagnosis systems should be able to detect their own potential misdiagnosis. This paper proposes a rule-based diagnostic validation algorithm using a previously developed two-stage diagnosis model in abnormal situations. We analyzed the diagnostic results of the sub-procedure stage when the first diagnostic results were inaccurate and derived a rule to filter the inconsistent sub-procedure diagnostic results, which may be inaccurate diagnoses. In a case study, two abnormality diagnosis models were built using gated recurrent units and long short-term memory cells, and consistency checks on the diagnostic results from both models were performed to detect any inconsistencies. Based on this, a re-diagnosis was performed to select the label of the second-best value in the first diagnosis, after which the diagnosis accuracy increased. That is, the model proposed in this study made it possible to detect diagnostic failures by the developed consistency check of the sub-procedure diagnostic results. The consistency check process has the advantage that the operator can review the results and increase the diagnosis success rate by performing additional re-diagnoses. The developed model is expected to have increased applicability as an operator support system in terms of selecting the appropriate AOPs and sub-procedures with re-diagnosis, thereby further increasing abnormal event diagnostic accuracy.

1. Introduction

1.1. Background

A nuclear power plant (NPP) is a large complex system with thousands of individual components. Each of them performs its own role, influences others, and is monitored with various parameters. The operators of the main control room diagnose the current plant state and predict the future plant state by referring to these parameters and all alarms. In all diagnoses, safety is considered a top priority [1], and to achieve high levels of safety, NPPs provide operating procedures that are suitable for certain situations. The operator selects the appropriate operating procedure by comparing it with the parameters and alarms indicating the current status. This allows the operator to make a diagnosis appropriate for each situation and take corrective action. The procedures that are the focus of this paper are the abnormal operating procedures (AOPs). The AOPs are procedures for responding to any abnormality and consist of several sub-procedures. Each sub-procedure represents different causes of the abnormality and has different entry conditions for each symptom and alarm. When an abnormal event occurs at an NPP, the current symptoms are compared to the entry conditions of each AOP and sub-procedures, and based on this the appropriate AOP is selected and corrective action is taken. At this point, if the situation becomes too serious to take actions through the AOP, or the situation worsens due to misdiagnosis or incorrect

* Corresponding author.
E-mail address: sjlee420@unist.ac.kr (S.J. Lee).

https://doi.org/10.1016/j.net.2022.05.031

action, the abnormal situation may become an emergency situation, causing the reactor to shut down and the plant to transit to a high standby state [2]. Therefore, operators are trained to select the appropriate AOPs and sub-procedures suitable for the situation within a short time.

In the case of an accident with a reactor trip, the necessary actions are taken in reference to the appropriate emergency operating procedure, which takes only a few minutes to identify as there are only seven emergency operating procedures in, for example, the Advanced Power Reactor 1400 (APR-1400) covered in this study. In contrast, the APR-1400 includes a total of 82 AOPs with 224 sub-procedures. Operators are trained to diagnose abnormal events, but it is difficult to compare the numerous parameters, alarms, etc. with more than 200 entry conditions, and this process takes a long time [3]. In addition, clear identification of an abnormal event may be difficult because some symptoms may be common to multiple AOPs and involve multiple alarms that can affect each other. Therefore, a diagnosis system that can support the operators is needed because confusion and human error in the diagnosis process may prevent the operator from selecting the appropriate AOP and taking the appropriate corrective action.

1.2. Related works

1.2.1. Related research

There has been a lot of research and effort over the past few decades to solve the problem of the difficulty in abnormal event diagnosis. In the beginning of these studies, methods of directly analyzing the parameters or patterns of the power plant were applied. Horiguchi et al. [4] employed the usage patterns of 49 power plant parameters to train an artificial neural network (ANN) to diagnose the causes of abnormalities. Lu et al. [5] introduced the group method of data handling modeling approach for system characterization, which works well in detecting the fault condition of devices during transient operations. Santosh et al. [6] diagnosed four transient initiating events using an ANN with a trained resilient-back propagation algorithm, and later developed a symptom-based diagnostic system to diagnose plant initiating events and detect deviations from normal operating conditions [7]. Serker et al. [8] used Elman's recurrent neural network (RNN) to detect bearing damage. Zio et al. [9] proposed a hierarchical structure to perform the classification of anomalies, where the data from known anomalies was used for one-class support vector machine training, while a multiclass support vector machine recognized the class to which transient data indicating anomalies belongs. Galbally et al. [10] suggested a system that can automatically classify transients using a dynamic time warping algorithm. Tolo et al. [11] proposed a combination of a set of ANN architectures through the use of Bayesian statistics for the detection and diagnosis of a loss of coolant accident. Despite the great progress, most models have been developed to diagnose emergency situations [12], and use only limited parameters depending on the specific situation without considering all parameters in the power plant. This is inadequate for diagnosing abnormal events with various causes and symptoms.

Recently, more advanced methodologies have been proposed by combining several machine learning algorithms or deep learning methods. Ayodeji et al. [13] performed fault diagnosis by combining principal component analysis (PCA) and recurrent neural network (RNN) algorithms. PCA plays a role in feature extraction and dimension reduction, and the developed model was applied to diagnose a total of five faults. Peng et al. [14] used 36 plant parameters to diagnose accidents and classified faults using a deep belief network. Furthermore, recently, models for diagnosing abnormalities using data-based methods with ANNs or convolutional neural networks have also been developed. Lee et al. [15] developed a convolutional neural network model to diagnose abnormality in around 10 different abnormal events. Kim et al. [16] also developed an abnormality diagnosis model that adopts a gated recurrent unit (GRU) and the PCA preprocessing method. In this work, the authors proposed a two-stage GRU model, that separately diagnoses AOPs and their sub-procedures. This model is the basis of the current study; details of the previously developed model are given below.

1.2.2. Preceding research

In the preceding research, a two-stage model was developed to diagnose an abnormal event by first selecting the appropriate AOP and then selecting the appropriate sub-procedure of the AOP [16]. Since the model training, generation, and diagnosis are conducted separately for the AOPs and sub-procedures, accuracy is improved by reducing the number of classes to be predicted by each model. As all 2829 parameters of the power plant simulator used for the model's data extraction are considered for abnormality diagnosis, a very large capacity is required, which makes pre-processing essential. In this case, the PCA method was applied to significantly reduce the dataset size while maintaining as much information as possible. Pre-processing succeeded in including more than 99% of the existing information using only 20 principal components. The training dataset was selected based on a total of 10 AOPs of the reference APR-1400. A GRU was used in the diagnostic model, and diagnostic accuracy was more than 99%. Fig. 1 depicts the overall process of the two-stage model. An AOP is selected by the main algorithm in the first stage, and the sub-procedure of the selected AOP is selected by the sub-algorithm in the second stage. In other words, it has advantages in accuracy and time because it trains the model and diagnoses the AOP and sub-procedure separately.

1.3. Purpose

Even though the abnormality diagnosis model of the preceding study achieved a high accuracy of about 99%, research is needed to further improve this accuracy. This is because NPPs are one of the representative examples of safety-critical systems [17], meaning that system failures may endanger human lives, cause significant economic damage, or cause extensive environmental damage. Accordingly, an extremely high level of safety is required and high reliability must be obtained in NPPs. However, data-driven models at present cannot always ensure accurate diagnosis because they cannot simulate all possible abnormal events, and thus the models' judgments need to be reviewed to improve the reliability of their diagnosis results.

For this purpose, a self-validation process that detects a false diagnosis by analyzing the judgment result of the model followed by a re-diagnosis to improve the diagnosis accuracy can be the one of the solutions. To select AOPs and sub-procedures in abnormal situations, this study builds two separate two-stage abnormality diagnosis models: one adopting GRUs, and one adopting long short-term memory (LSTM) cells. Using the characteristics of the models, we propose a rule-based diagnostic validation called the consistency check algorithm that is used in a re-diagnosis process. The consistency of the sub-procedure diagnostic results is checked to filter the inconsistent results, which may be an incorrect diagnosis. Based on these cases, re-diagnosis selecting the second-best AOP is performed, and accuracy is improved. In this way, this study aims to increase the applicability of the developed model as an operator support system by increasing diagnosis performance through detecting the model's own misdiagnosis and re-diagnosing.

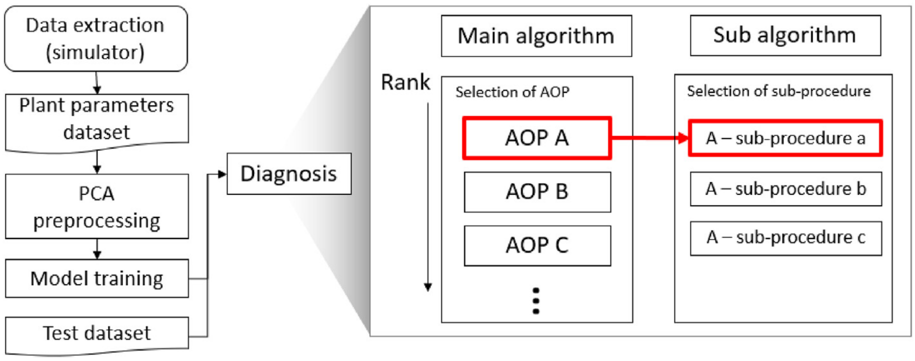

Fig. 1. Diagnosis process of the two-stage diagnosis model using GRU [16].

2. Methodology

2.1. Framework

To enable self-validation and re-diagnosis by the developed model to detect misdiagnosis, we propose the following method. Fig. 2 presents a framework that schematizes the overall process covering data generation for the base diagnostic model, artificial intelligence model generation, AOP and sub-procedure selection, self-validation through a consistency check method, and re-diagnosis. The specific steps are as follows.

- Step 1: Create a base diagnostic model referring to the two-stage model proposed in the preceding study [16]. Use this model to make the first diagnosis selecting the AOP and sub-procedure.
- Step 2: Analyze the diagnostic results using the proposed *consistency check* method. This method serves to filter out possible cases of misdiagnosis by analyzing the results of the sub-procedure selection. For this, a filtering method using three factors is applied.
- Step 3: Re-diagnose the cases filtered as *inconsistent* with a high probability of false diagnosis in the previous step.
- Step 4: Go back to step 2 and conduct a consistency check on the results of the re-diagnosis, namely the new sub-procedure selection. Repeat this process until there are no inconsistent results.

In short, the goal is to achieve high reliability of successful diagnosis by performing re-diagnosis on the results that may be misdiagnosed. While the above steps may be repeated as many times as necessary, in this study, we achieved improved diagnosis performance with only two diagnosis and consistency check processes (see Section 4).

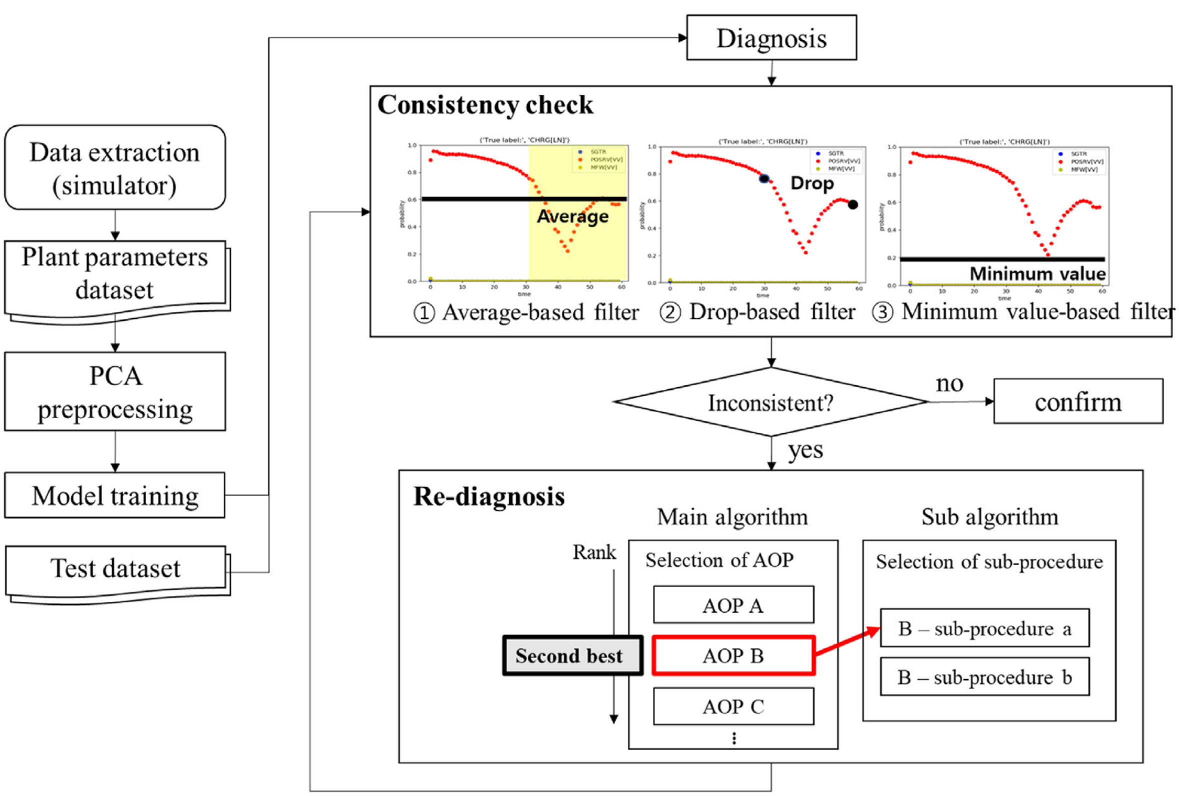

Fig. 2. Framework of the current study.

2.2. Consistency check

The first part of our proposed method is a consistency check. In the previous section, it was discussed that the AOP and the sub-procedure are separately diagnosed as a characteristic of the base diagnostic model. The diagnostic process for two cases is shown in Fig. 3 as an example. After diagnosis by applying the diagnostic model to test dataset, the results of the selected AOP and sub-procedure can be expressed as a probability graph over diagnosis time. In the case of the first example, the diagnosis of the AOP and sub-procedure was successful, with the sub-procedure graph showing a *consistent* trend in which the probability value of one label is close to 1. However, failure to diagnose shows a different trend in the results. In the case of misdiagnosis, unlike the above case, the sub-procedure diagnostic probability graph shows an *inconsistent* trend that decreases rapidly rather than a consistent graph close to 1.

In other words, *consistent* means that one label in the graph can be inferred as the correct answer, that is, a result with a high probability of one label is obtained. Conversely, *inconsistent* is a concept proposed to indicate a case in which the diagnosis result cannot be predicted or confirmed. This concept was applied because if a misdiagnosis by the main algorithm led to the selection of the wrong AOP, the sub-algorithm also selects the wrong sub-procedures. That is, misdiagnosis by the sub-algorithm is inevitable, and an inconsistent graph appears due to the contradiction of having to select one among all wrong options.

Therefore, self-validation that filters any *inconsistent* cases by analyzing the sub-procedure diagnosis results is essential, and through this process, we are able to get an opportunity for re-diagnosis.

In order to find the optimal means to filter the *inconsistent* results, we first analyzed the probability graph according to the diagnosis time of the sub-algorithm when the diagnosis failed. As a result of examining the resulting graph of misdiagnosis, we discovered trends such as multiple sub-procedure appearances, a sharp decrease in probability, a low average value of probability, or no sub-procedure appearances. Referring to the graphs shown in Fig. 4, the first example is a case where dataset from SGTL was misdiagnosed as CHRG with several sub-procedures appearing, and the second example is a case where dataset from RMW was misdiagnosed as CWS with a rapid decrease in probability. In the third example, dataset from POSRV was misdiagnosed as RMW, showing a trend with no sub-procedure appearance.

It would be difficult to judge the graphs of these inconsistent trends simply by looking at them. Therefore, in order to find the reference points and conditions that can be used to filter the inconsistent results, sensitivity studies were conducted by setting several factors, namely an average-based filter, a drop-based filter, and a minimum value-based filter. First, it can be seen that multiple sub-procedure appearances showed a very low probability value when the label of the maximum value was changed, and also that the average value was likewise low. In the case of sharp decrease, the average value after 30 s was low and the trend of decrease was confirmed. It can be seen that the low average trend also appears in most inconsistent cases, and the average and minimum values are also very low in the no sub-procedure appearance trend. Accordingly, the average-based method can filter low average value and no sub-procedure appearance trends, and the minimum value-based method can filter multiple sub-procedures, and no sub-procedure. The drop-based method can filter the sharp decrease trend. These filtering methods can be applied to several trend in common, not to the standards created in response to one trend each. By applying this, the threshold of each filtering method is adjusted little by little to find the optimal condition. Fig. 5 shows examples of these filtering methods. The first factor checks whether the average probability value after 30 s is lower than a pre-determined threshold. The second factor compares the early and the later sections or the middle and the later sections in 1 min to check whether the probability decreases by more than 0.1. The third factor checks whether the minimum value of the probability is lower than a pre-determined threshold. It is necessary to find the optimal combination while finely adjusting the threshold conditions of these factors, because the more cases of failed diagnosis filtered as *inconsistent* and the more cases of successful diagnosis filtered as *consistent*, the better the filtering performance. In other words, a consistency check is performed on the sub-procedure diagnosis results obtained by the base diagnostic model, which makes it possible to identify a misdiagnosis from an inconsistent result, giving an opportunity for re-diagnosis.

2.3. Re-diagnosis

The second part of the proposed method is re-diagnosis. Re-diagnosis is required any results of the consistency check are filtered as *inconsistent*. When moving to re-diagnosis, the AOP with the highest probability is excluded from the procedure selection process of the main algorithm of the previous first diagnosis. In

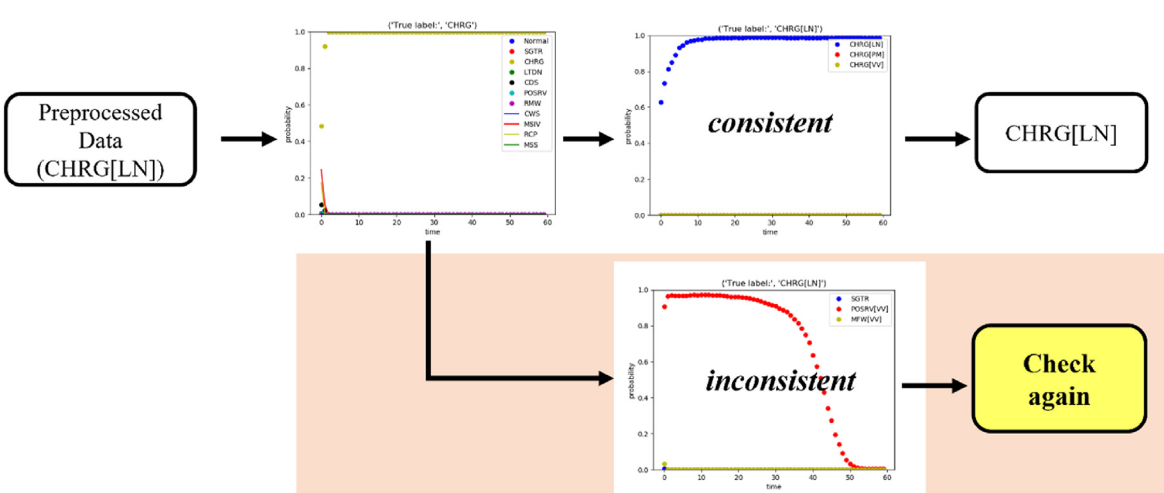

Fig. 3. Sub-procedure diagnosis result graph trends.

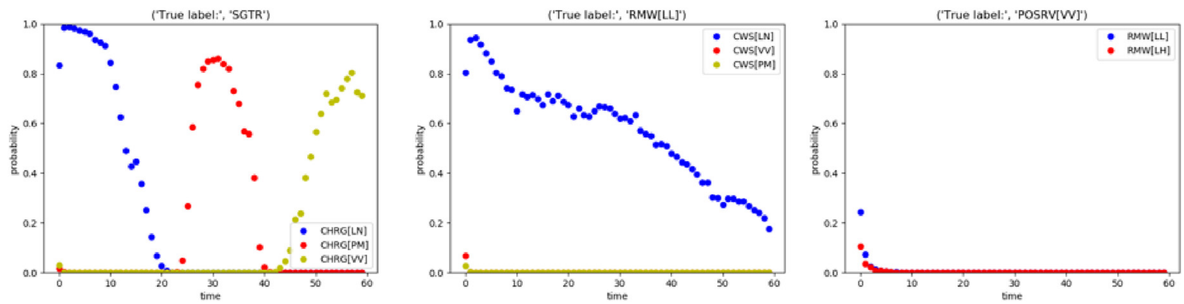

Fig. 4. Example of *inconsistent* sub-procedure diagnostic results.

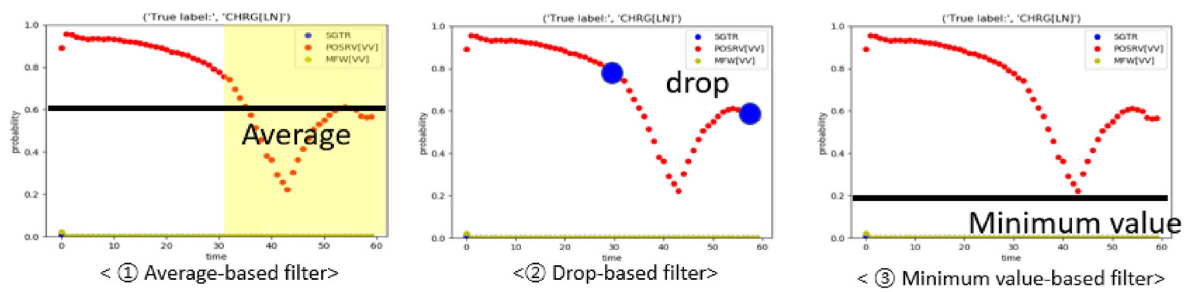

Fig. 5. Example of three filtering methods in the sub-procedure diagnostic result graphs.

more detail, re-diagnosis is performed by arbitrarily assigning a value less than 0 to the AOP misdiagnosed in the first diagnosis and then selecting the second-best procedure with the highest value among the remaining AOPs. The sub-procedure is also selected from among the sub-procedures corresponding to the second-best AOP. This process is illustrated in Fig. 6.

3. Experimental setting

3.1. Data configuration

Training datasets are needed to create the base abnormality diagnosis model. In addition, large amounts of datasets are needed to compare with the results of the previous study and for more realistic analysis. However, due to the lack of records of abnormal events in actual NPPs, we used a simulator to extract data from scenarios of abnormal events that could actually occur. In this study, a 3KEYMASTER full-scope simulator made by Western Corporation Services was used [18], as shown in Fig. 7. As a simulator of a 1400 Mwe pressurized water reactor, it has the advantage of being able to apply abnormal scenarios of the APR-1400 referenced in this study. In particular, since the simulator covers almost all the components and functions of an actual NPP system, it shows symptoms similar to those in real abnormal events in actual NPPs. A pre-written scenario script was applied to inject the intended abnormal symptoms, after which changes in a total of 2829 parameters for 1 min were recorded and saved.

The scenarios were selected based on the cases in which abnormal symptoms were sufficiently realized through the 3KEYMASTER simulator. 19 AOPs were selected by adding scenarios judged to be implementable in the simulator to 10 AOPs selected in the preceding study [16]. To obtain the dataset, we first set the initial conditions of the simulator to 100% power generation without any events. The scenario suitable for each AOP and the sub-procedure was manually implemented and the corresponding malfunctions, indicating component failures such as valve opening/closing failure, or pump stop, were injected.

In order to diversify the scenarios, 300 datasets were generated for each sub-procedure by finely dividing the area between the minimum and maximum values of the function and combining the time of the change of the function value and the switching conditions. The dataset selected all 2829 parameters that could indicate a

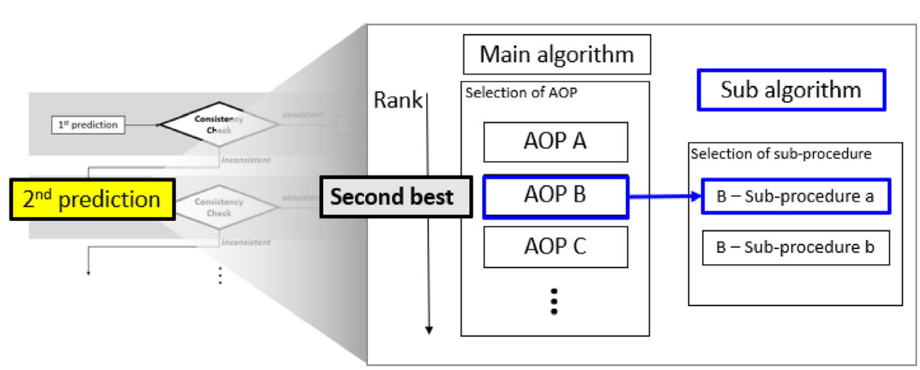

Fig. 6. Process of re-diagnosis selecting the second-best AOP.

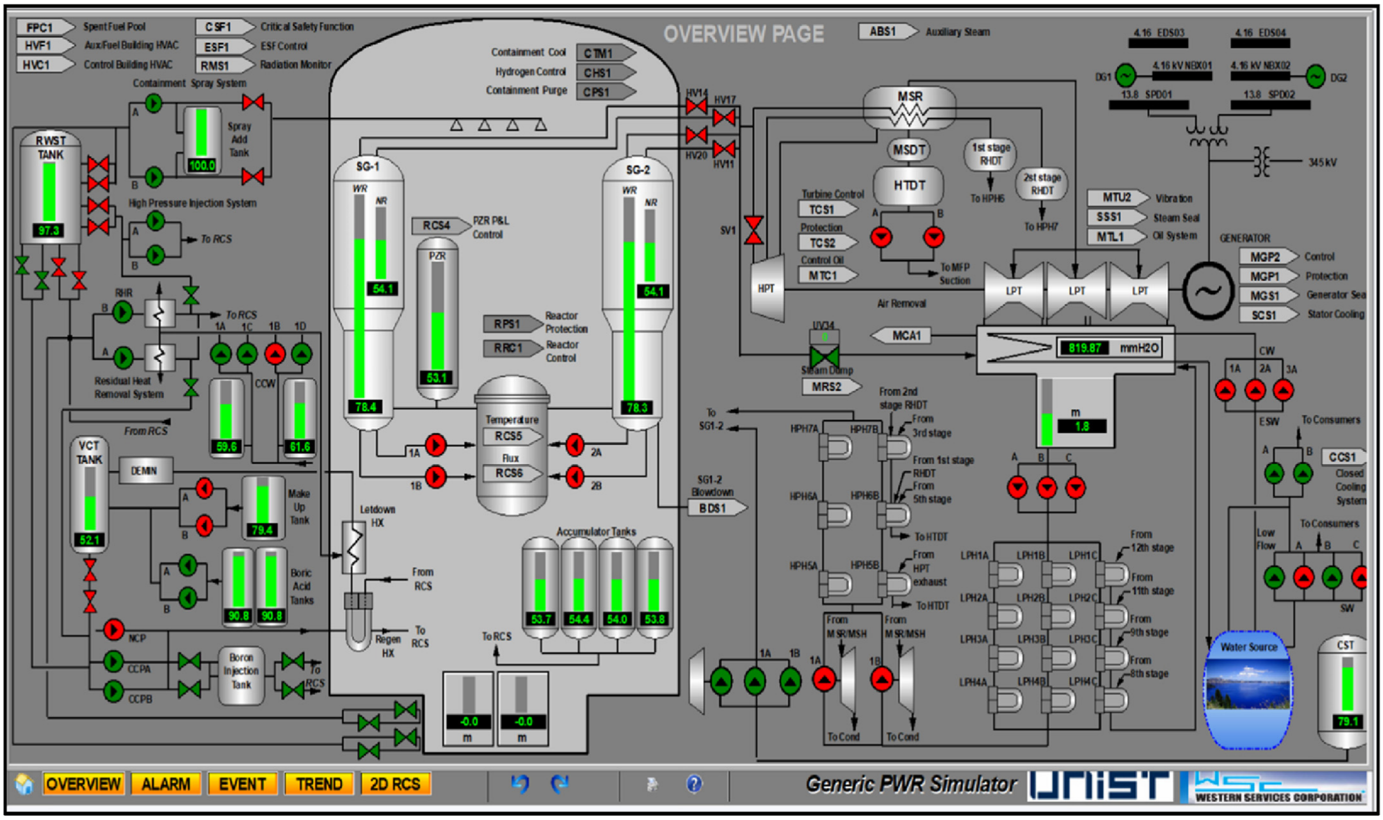

Fig. 7. Screenshot of the 3KEYMASTER NPP simulator.

change in the state of the simulator. The simulator was executed by inputting the preset scenario files. 2829 parameters accumulated state values once per second to obtain 2829 × 60 variables per scenario.

In the simulations, cases in which an alarm was generated but the reactor tripped within 1 min and cases in which there was no alarm or variable change for 1 min were excluded.

Unlike the previous study that selected 10 AOPs [16], the current study selected 19 AOPs because more results are required to analyze cases of misdiagnosis. The 19 total AOPs with 34 sub-procedures selected and simulated from abnormal scenarios in this work are listed in Table 1.

The 19 total abnormal events corresponding to the selected AOPs mainly comprise events that are important in terms of probability of occurrence or outcome. These abnormal events cover the full range of NPP primary, secondary, and support and safety systems. In addition, the relevance of each target system and the number of sub-procedures were also considered. Out of the total 34 sub-procedures, 32 procedures generated 300 datasets for each procedure, and 2 procedures generated 30 and 260 datasets, respectively, depending on the simulator situation. In addition, 300 datasets of a normal state were also generated, resulting in a total of 10,190 datasets.

3.2. Model training

Based on the extracted training datasets, it is necessary to create a model for diagnosing NPP abnormalities. For model training, a total of 20 labels were considered by combining the 19 AOPs and the normal state in the main algorithm, and the sub-algorithms were trained separately for each abnormality and the corresponding sub-procedures were considered in 19 sub-algorithms. Of the total 10,190 datasets, 1978 datasets, or 20% of the 9890 datasets corresponding to the abnormality dataset, were excluded to be used as test datasets. The remainder was used as training datasets for model training. That is, the dataset can be largely divided into training and test datasets. For example, out of 300 datasets from one sub-procedure, the 5th, 10th, 15th … 300th datasets are set as the test dataset, the remaining 240 datasets will be the training dataset.

It may be said that an RNN is a suitable algorithm for processing the data of an NPP simulator having the characteristics of time series data. However, it has a disadvantage of deteriorating performance as the time delay increases in the backpropagation process. Therefore, we went through the process of creating models separately using LSTMs and GRUs, algorithms shown to solve the long-term dependency problem of existing RNNs [19].

In the training process, K-fold cross-validation was performed to evaluate the performance of the developed model. Setting K = 5, this method randomly divides the training datasets into five groups, trains the model using the datasets from four groups as training datasets, and then conducts testing using the remaining group. A total of five validations were made by repeating this process while varying the validation dataset group [20,21]. In this process, k-fold cross-validation is performed to create five models from the same dataset. The 1/n of training datasets may be validation datasets arbitrarily according to the number of repetitions (n) set for the k-fold cross-validation. This validation dataset is not fixed, but is randomly changed by repeating cross-validation of the model.

When we measured the training accuracy of the model following the above training process, the average accuracy was about 99% in the main algorithm and 100% in the sub-algorithm. The accuracy of the main algorithm can be seen in more detail in

Table 1
Selected AOPs and sub-procedures.

AOPs	Sub-procedures[a]
Steam generator tube leakage (SGTL)	SGTL
Charging water system abnormality (CHRG)	CHRG[PM], CHRG[VV], CHRG[LN]
Letdown water system abnormality (LTDN)	LTDN[LN], LTDN[VV]
Condenser vacuum abnormality (CDS)	CDS
Pilot-operated safety relief valve leakage (POSRV)	POSRV[VV]
Reactor makeup water tank valve abnormality (RMW)	RMW[LL], RMW[LH]
Circulating water system abnormality (CWS)	CWS[LN], CWS[VV], CWS[PM]
Main steam isolation valve abnormality (MSIV)	MSIV
Reactor coolant pump abnormality (RCP)	RCP[LC], RCP[SD], RCP[SL]
Main steam system abnormality (MSS)	MSS[VV], MSS[LN]
Pressurizer pressure low abnormality (PZR)	PZR[VV], PZR[AV]
Low pressure feedwater heater level high abnormality (LFH)	LFH[VV], LFH[TB]
High pressure feedwater heater level high abnormality (HFH)	HFH[VV], HFH[LN], HFH[TB]
Main feedwater pump recirculation valve abnormality (MFW)	MFW[VV]
High pressure turbine control valve abnormality (TCS)	TCS[VV]
Turbine generator building closed cooling water system abnormality (CCS)	CCS[PM]
Component cooling water system abnormality (CCW)	CCW[SL], CCW[XL]
Spent fuel pool cooling abnormality (FPC)	FPC[PM], FPC[VV]
Turbine control oil system abnormality (MTC)	MTC[PM]

[a] PM: pump trip, VV: valve abnormality, LN: line (tube) leakage, LL: low level, LH: high level, LC: CCW loss, SD: seal damage, SL: seal injection water loss, AV: auxiliary valve abnormality, TB: tube rupture, XL: heat exchanger leakage.

Table 2. The sub-algorithm showed 100% accuracy in all 19 algorithms. Very high accuracy is inevitably obtained because the sub-algorithm selects one among only two or three sub-procedures of each AOP. At this time, in the case of an AOP with only one sub-procedure, there is only one option for sub-procedure selection, so even if the AOP is misdiagnosed, the sub-procedure diagnosis results are bound to be *consistent*. In this case, a classification model cannot be created. In addition, the more various patterns are learned, the more the accuracy can be increased. Therefore, two or three grouped sub-procedures make one sub-algorithm and share it to diagnose sub-procedures. Of the total 19 AOPs, SGTL, POSRV, and MFW share one sub-algorithm, CDS, MSIV, and TCS share one, and MTC, and CCS share one, so this gives a total of 14 different sub-algorithms.

Comparing this in a diagnostic graph, it can be seen that there is a significant difference. The left panel of Fig. 8 shows a *consistent* result because there is only one option for the sub-procedure diagnosis despite the misdiagnosis by the main algorithm. Conversely, the right panel shows *inconsistent* results according to the misdiagnosis because several sub-procedures were grouped and trained together in the sub-algorithm. Therefore, since the sub-algorithm selects one among the two or three sub-procedures, if the result of this selection is *inconsistent*, the previous AOP diagnosis is highly likely to be wrong.

3.3. Consistency check criteria setting

As mentioned in Section 2.2, most successful diagnoses show a consistent trend, while a misdiagnosis shows an inconsistent trend. Such *inconsistent* results were classified into three types, as shown in Table 3. However, there were cases in which there was an opposite tendency in the early stage of this work. The criteria and thresholds of consistency to minimize these opposite cases were empirically selected through several experiments. Initial tests and sensitivity studies were conducted to establish conditions and criteria for optimizing the performance of the consistency check.

In the sensitivity studies, we set different factors that filter out the *inconsistent* results. The three factors were an average-based filter [a)], a drop-based filter [b)], and a minimum value-based filter [c)]. By finely adjusting the conditions of these factors, we could find their optimal combination. Consistency checks performed while changing the criteria and threshold of each factor revealed that the number of cases detected as *inconsistent* also changed as the threshold of each factor changed. Therefore, the optimal conditions were selected according to adjusting the combinations of the numerical values of each factor; the results are shown in Table 4. The figure in (a) is based on whether the average value in the second half is below the corresponding numerical value. The factor (b) checks whether it decreases by comparing the early and the later sections or the middle and the later sections. In this case, the probability values of each section were compared by setting the early section to 10 s, the middle section to 30 s, and the later section to 60 s. Therefore, it can be understood that if 10 and 60 are presented in Table 4, the early and later sections are compared, and if 30 and 60 are presented, the middle and later sections are compared. The figure in (c) is to see if the minimum value is less than or equal to the numerical value in the trend of the dataset for 60 s.

The more cases in which failed diagnoses are filtered as *inconsistent* and the more cases in which successful diagnoses are filtered as *consistent*, the better the filtering performance is. In the GRU diagnostic model, when the threshold of the average-based filter was 0.5 and a minimum value-based filter was 0.1, that is, conditions 1 and 3 were the best combination criteria. In the LSTM diagnostic model, conditions 3, 4, 7, and 8 were the best criteria. Accordingly, it was found that the best condition common to both diagnostic models was condition 3, and therefore the consistency check was performed based on condition 3.

Table 2
Cross-validation average accuracy of the first stage.

Main algorithm	#1	#2	#3	#4	#5	Average
GRU	0.9977	0.9982	1.0000	0.9542	1.0000	0.99
LSTM	0.9994	0.9994	0.9994	0.9918	0.9994	0.9978

4. Results

4.1. Base model diagnosis

First, diagnostic results were obtained from the base model, as shown in Table 5. More specifically, the results are from a total of 1978 test datasets applied to two separate two-stage models based

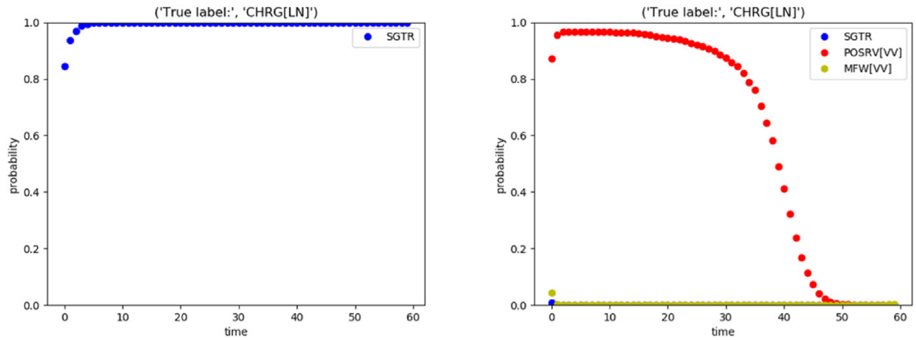

Fig. 8. Comparison of misdiagnosis results between sub-algorithms with one sub-procedure (left) and several sub-procedures (right).

Table 3
Classification into three *inconsistent* types of sub-procedures.

Type	GRU	LSTM
Multiple sub-procedures	2	1
Rapid reduction	3	2
No sub-procedures	5	5

Table 5
Diagnostic results of the base model.

	# of cases	Success	Failure	Misdiagnosis rate	Diagnosis accuracy
GRU	1978	1945	33	1.668%	98.332%
LSTM	1978	1962	16	0.809%	99.191%

on GRU and LSTM. In Table 5 and the following tables, the misdiagnosis rate and the diagnosis accuracy are defined as follows:

- Misdiagnosis rate: $\frac{\textit{Diagnosis failure cases}}{\textit{All cases}}$
- Diagnosis accuracy: $\frac{\textit{Diagnosis success cases}}{\textit{All cases}}$

In the GRU model, there were 33 diagnosis failures, and the misdiagnosis rate was 1.668%. In the LSTM model, there were 16 diagnosis failures, and the misdiagnosis rate was 0.809%.

4.2. Consistency check

By performing a consistency check, the results could be organized by dividing them into *consistent* and *inconsistent*. When the consistency check was performed in the current case study based on the diagnosis results of the base model, 7 cases in the GRU model and 13 cases in the LSTM model were filtered as *inconsistent*; this is schematically illustrated in Fig. 9 and Fig. 10, respectively. Here, it can be seen that the misdiagnosis rate was reduced after filtering (excluding) the *inconsistent* labels. Results are summarized in Table 6.

It is necessary to explain in more detail the concepts of consistent or inconsistent and diagnosis success or failure shown in Table 6. As proposed in this work, the consistency check is important to filter out uncertain results in situations where the operator does not know whether the diagnostic results are accurate. In the case of a consistent result, since it is a confirmed result, whether the diagnosis is correct or not can be expressed as diagnosis success or diagnosis failure. But in the case of an inconsistent result, re-diagnosis is required, and therefore it can be considered as an unknown result rather than diagnosis failure.

Accordingly, looking at the results after the consistency check, the misdiagnosis rate decreased from 1.668% to 1.314% because 7 cases out of 33 diagnosis failures were excluded as inconsistent in the GRU model. In the LSTM model, the misdiagnosis rate decreased from 0.809% to 0.151% because 13 cases out of 16 diagnosis failures were excluded as inconsistent. In other words, while the diagnosis success rate remained the same, the misdiagnosis rate was lowered from the creation of the inconsistent label.

4.3. Re-diagnosis

Re-diagnosis is necessary for the results filtered as inconsistent in the previous consistency check. The process from the first diagnosis results through the consistency check to re-diagnosis can be simply illustrated for the GRU and LSTM models as shown in top panels of Figs. 9 and 10, respectively. As shown in Figs. 9 and 10 and Table 6, there were 7 cases of inconsistent in the GRU model and 13

Table 4
Sensitivity study of filtering factors.

—	Factor				Inconsistent (GRU)		Inconsistent (LSTM)	
	(a)	(b)		(c)	Diagnosis success	Diagnosis failure	Diagnosis success	Diagnosis failure
1	0.5	10	60	0.1	0	9	3	8
2	0.5	10	60	0.2	1	9	3	8
3	0.5	30	60	0.1	0	9	1	8
4	0.5	30	60	0.2	1	9	1	8
5	0.6	10	60	0.1	3	9	3	8
6	0.6	10	60	0.2	3	9	3	8
7	0.6	30	60	0.1	3	9	1	8
8	0.6	30	60	0.2	3	9	1	8
9	0.7	10	60	0.1	24	9	47	8
10	0.7	10	60	0.2	24	9	47	8
11	0.7	30	60	0.1	23	9	20	8
12	0.7	30	60	0.2	23	9	20	8

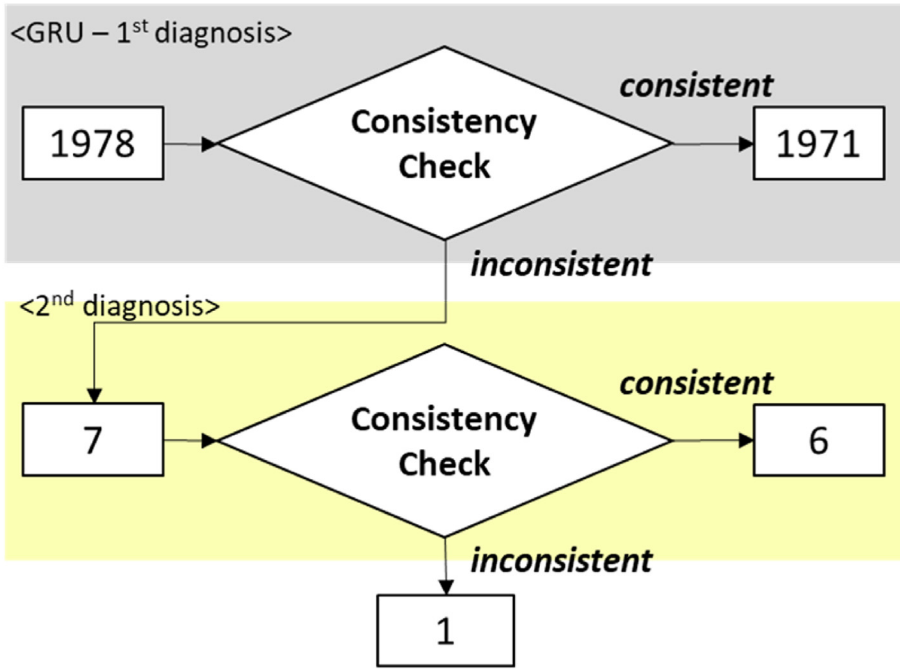

Fig. 9. Diagram of the consistency check and re-diagnosis results of the GRU diagnostic model.

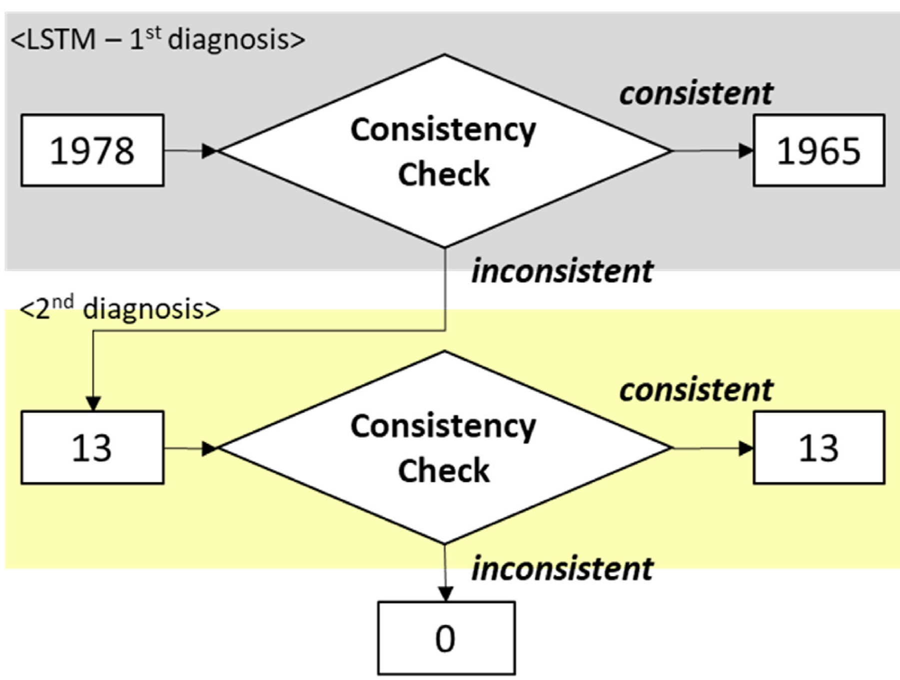

Fig. 10. Diagram of the consistency check and re-diagnosis results of the LSTM diagnostic model.

Table 6
Consistency check results.

	# of cases	Consistent		Inconsistent	Misdiagnosis rate	Diagnosis accuracy
		Success	Failure			
GRU	1978	1945	26	7	1.314%	98.332%
LSTM	1978	1962	3	13	0.151%	99.191%

Table 7
Re-diagnosis results.

	# of cases	Consistent		Inconsistent	Misdiagnosis rate	Diagnosis accuracy
		Success	Failure			
GRU	1978	1951	26	1	1.314%	98.635%
LSTM	1978	1975	3	0	0.151%	99.848%

cases in the LSTM model. When a re-diagnosis of these cases was conducted, 6 cases were successfully re-diagnosed and shown as *consistent* in the GRU model, and all 13 cases were successfully re-diagnosed and shown as *consistent* in the LSTM model. This is expressed in Table 7. As the number of cases of diagnosis success increased through re-diagnosis, the diagnosis accuracy was also improved in both models.

5. Discussion

We created a diagnostic model using GRU and LSTM, and we were able to confirm the diagnostic prediction results using the time series data of the nuclear power plant simulator. Through PCA preprocessing, it was possible to preserve 99% of the original data information using only 20 PCs with 2829 parameters. This prevented many parameters from complicating the diagnostic algorithms and also reduced computational time. Subsequently, it was found that there was a difference in the case of diagnosis success and failure by graphing the diagnostic prediction values for 1 min. If misdiagnosed in the first stage, inconsistency is clearly visible in the graph of second stage because the sub-algorithm of other AOP misdiagnosed is brought and diagnosed. Organizing this into several cases, various methods have been proposed, tested, and many attempts have been made to find more accurate filtering conditions. Based on this, it was confirmed that the filtered inconsistent cases found the correct answer through re-diagnosis.

We compared the diagnosis results of the base model, the results of a consistency check of the first diagnosis, and the results of one re-diagnosis in section 4. Through the self-validation process that filters misdiagnoses as *inconsistent* through the consistency check, the misdiagnosis rate could be reduced, and based on this, re-diagnosis was conducted which increased the overall diagnosis accuracy.

After that, future efforts are needed to increase the model performance and apply it in actual plant situation. In this study, model construction and diagnosis were performed by applying only 19 AOPs, which still insufficiently reflect the actual 82 AOPs of the reference plant. In addition, since the experiments were conducted with data obtained through a simulator on account of the lack of records of abnormal events in actual power plants, it is necessary to solve the noise problem, which refers to discrepancies between real and simulated plant data.

Moreover, this study considered only three filtering methods in the consistency check. It will be possible to further improve the accuracy of the model by applying various alternative filtering methods. In addition to comparing the trends with the thresholds in the consistency check, applying another algorithm in the filtering process can also be attempted.

6. Conclusion

In this study, a method was proposed for improving the accuracy of abnormality diagnosis in an NPP through adding self-validation and re-diagnosis features to an abnormal event diagnostic model. NPPs have numerous parameters and alarms that change over time, which can be monitored and used to diagnose the plant condition. Abnormal events occurring in NPPs have a lower risk than ones in emergency situation, but since there are dozens of possible abnormalities, it is quite difficult to consider them all in a short time for accurate diagnosis, where operators should compare parameters and alarms to select the appropriate operating procedures to make a situation-specific diagnosis and take corrective action. In this light, if a misdiagnosis or wrong action is taken, the situation may worsen, resulting in a shutdown of the reactor and transition to a hot standby state. Accordingly, a neural network model for diagnosing abnormalities has been previously developed, but since accurate diagnosis cannot always be guaranteed, the diagnostic model should be able to detect and review possible misdiagnosis on its own.

The proposed method comprises a review of the initial diagnostic results, a consistency check, and self-validation and re-diagnosis processes. First, the appropriate AOP and sub-procedure are diagnosed using the two-stage model, and the results are analyzed to find the optimal filtering method for the consistency check. Diagnosis accuracy can be improved by performing re-diagnosis on the *inconsistent* cases with a high possibility of misdiagnosis obtained through the consistency check. While the levels of increased accuracy may seem numerically small, any improvement is highly significant due to the nature of NPPs requiring extremely high reliability.

Finally, the possibility is open to develop the model into a recursive structure by applying re-diagnosis and termination criteria. Through this, a recursive model can realize very high accuracy, which further increases its applicability as an operator support system in real plants.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety (KoFONS) using a financial resource granted by the Nuclear Safety and Security Commission (NSSC) of the Republic of Korea (No. 2106007) and also by Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government (MOTIE) (No. 20211510100020, Development of operation supporting technology based on artificial intelligence for nuclear power plant start up and shutdown operation).

References

[1] M.J. Embrechts, S. Benedek, Hybrid identification of nuclear power plant transients with artificial neural networks, IEEE Trans. Ind. Electron. 51 (3) (2004) 686–693.
[2] S. Lee, P. Seong, A dynamic neural network based accident diagnosis advisory system for nuclear power plants, Prog. Nucl. Energy 46 (3–4) (2005) 268–281.
[3] J. Park, W. Jung, A systematic framework to investigate the coverage of abnormal operating procedures in nuclear power plants, Reliab. Eng. Syst. Saf. 138 (2015) 21–30.

[4] M. Horiguchi, N. Fukawa, K. Nishimura, Development of nuclear power plant diagnosis technique using neural networks, in: Proceedings of theFirst International Forum on Applications of Neural Networks to Power Systems, IEEE, 1991, pp. 279–282.
[5] B. Lu, B.R. Upadhyaya, "Monitoring and fault diagnosis fo the steam generator system of a nuclear power plant using data-driven modeling and residual space analysis," Ann. Nucl. Energy, Volume 32, Issue 9, pp.897-912.
[6] T.V. Santosh, G. Vinod, R.K. Saraf, A.K. Ghosh, H.S. Kushwaha, Application of artificial neural networks to nuclear power plant transient diagnosis, Reliab. Eng. Syst. Saf. 92 (10) (2007) 1468–1472.
[7] T.V. Santosh, A. Srivastava, V.V.S. Sanyasi Rao, A.K. Ghosh, H.S. Kushwaha, Diagnostic system for identification of accident scenarios in nuclear power plants using artificial neural networks, Reliab. Eng. Syst. Saf. 94 (3) (2009) 759–762.
[8] S. Şeker, E. Ayaz, E. Türkcan, Elman's recurrent neural network applications to condition monitoring in nuclear power plant and rotating machinery, Eng. Appl. Artif. Intell. 16 (2003). Issues 7–8.
[9] M. Claudio, S. Rocco, Enrico Zio, A support vector machine integrated system for the classification of operation anomalies in nuclear components and systems, Reliab. Eng. Syst. Saf. 92 (5) (2007) 593–600.
[10] J. Galbally, D. Galbally, A pattern recognition approach based on DTW for automatic transient identification in nuclear power plants, Ann. Nucl. Energy 81 (2015) 287–300.
[11] Silvia Toloa, Xiange Tian, Nils Bausch, Victor Becerra, T.V. Santhosh, G. Vinod, Edoardo Patelli, Robust on-line diagnosis tool for the early accident detection in nuclear power plants, Reliab. Eng. Syst. Saf. 186 (2019) 110–119.
[12] Jianping Ma, Jin Jiang, Applications of fault detection and diagnosis methods in nuclear powerplants: a review, Prog. Nucl. Energy 53 (2011) 255–266.
[13] A. Ayodeji, Y. Liu, H. Xia, Knowledge base operator support system for nuclear power plant fault diagnosis, Prog. Nucl. Energy 105 (2018) 42–50.
[14] B.S. Peng, H. Xia, Y.K. Liu, B. Yang, D. Guo, S.M. Zhu, Research on intelligent fault diagnosis method for nuclear power plant based on correlation analysis and deep belief network, Prog. Nucl. Energy 108 (2018) 419–427.
[15] Gyumin Lee, Seung Jun Lee, Changyong Lee, A convolutional neural network model for abnormality diagnosis in a nuclear power plant, Appl. Soft Comput. 99 (2021), 106874.
[16] J.M. Kim, G. Lee, C. Lee, S.J. Lee, Abnormality diagnosis model for nuclear power plants using two-stage gated recurrent units, Nucl. Eng. Technol. 52 (9) (2020) 2009–2016.
[17] John C. Knight, Safety Critical Systems: Challenges and Directions", ICSE '02: Proceedings of the 24th International Conference on Software Engineering, 2002, pp. 547–550.
[18] 3KEYMASTER Simulator 2013, Western Service Corporation, Frederick, MD, USA.
[19] Rui Fu, Zuo Zhang, Li Li, Using LSTM and GRU Neural Network Methods for Traffic Flow Prediction", 31st Youth Academic Annual Conference of Chinese Association of Automation, YAC), 2016.
[20] Y. Bengio, Y. Grandvalet, No unbiased estimator of the variance of k-fold cross-validation, J. Mach. Learn. Res. 5 (2004) 1089–1105.
[21] J.M. Kim, S.J. Lee, Framework of Two-Level Operation Module for Autonomous System of Nuclear Power Plants during Startup and Shutdown Operation", KNS-2019, Korean Nuclear Society, 2019.