Nuclear Engineering and Technology 56 (2024) 558–567

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

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

Original Article

Approach to diagnosing multiple abnormal events with single-event training data

Ji Hyeon Shin [a], Seung Gyu Cho [a], Seo Ryong Koo [b], Seung Jun Lee [a,*]

[a] Department of Nuclear Engineering, Ulsan National Institute of Science and Technology, 50, UNIST-gil, Ulsan, 44919, Republic of Korea
[b] Korea Atomic Energy Research Institute, 111, Daedeok-daero 989 beon-gil, Yuseong-gu, Daejeon, 34057, Republic of Korea

ARTICLE INFO

Keywords:
Nuclear power plant
Abnormal event diagnosis
Multi-label classification
Feature selection

ABSTRACT

Diagnostic support systems are being researched to assist operators in identifying and responding to abnormal events in a nuclear power plant. Most studies to date have considered single abnormal events only, for which it is relatively straightforward to obtain data to train the deep learning model of the diagnostic support system. However, cases in which multiple abnormal events occur must also be considered, for which obtaining training data becomes difficult due to the large number of combinations of possible abnormal events. This study proposes an approach to maintain diagnostic performance for multiple abnormal events by training a deep learning model with data on single abnormal events only. The proposed approach is applied to an existing algorithm that can perform feature selection and multi-label classification. We choose an extremely randomized trees classifier to select dedicated monitoring parameters for target abnormal events. In diagnosing each event occurrence independently, two-channel convolutional neural networks are employed as sub-models. The algorithm was tested in a case study with various scenarios, including single and multiple abnormal events. Results demonstrated that the proposed approach maintained diagnostic performance for 15 single abnormal events and significantly improved performance for 105 multiple abnormal events compared to the base model.

1. Introduction

Each system in a nuclear power plant (NPP) consists of hundreds of components that are required to perform its given function. A malfunction or abnormality occurring in any of these components can cause the related system to fail, and if the abnormal event is exacerbated, the reactor may shut down and result in an emergency situation. As emergency situations in NPPs can create economic loss by halting electricity production and also threaten safety, it is important that the operators in the main control room mitigate abnormal events to avoid such outcomes. Operators must recognize an abnormal event within a given time based on information from key monitoring parameters and alarms and follow appropriate measures given in abnormal operating procedures [1]. These tasks mostly rely on their experience and skill. The abnormal event diagnostic process can be highly burdensome, which makes this process directly related to NPP safety. Therefore, developing a model with a diagnosis function as part of an operator support system can be helpful for improving NPP safety.

Abnormal events can occur in many various ways in actual NPPs, including not only the well-studied single events but also cases with multiple occurrences, or multi-abnormal events, which are more complex to diagnose. As an example of this complexity, multiple failures occurring in related systems such as the letdown line and the charging line can produce similar monitoring parameter changes. In addition, multiple failures simultaneously occurring in unrelated systems can produce complex monitoring parameter changes. That is, parameter changes in multi-abnormal events may differ from the alarms and symptoms described in the related abnormal operating procedures. These situations require a high-level diagnosis ability of operators. Accordingly, for an operator support system to perform its intended role in actual NPPs, diagnosis models have to maintain stable performance in situations with multi-abnormal events, despite their low probability.

Deep learning models applied to classification problems in many industries need sufficient training data for each target label to achieve high performance in classification [2]. In the case of the nuclear field, the target labels correspond to the target abnormal events, for which event diagnosis models likewise need to acquire sufficient data to achieve high diagnostic performance. Therefore, it could be thought that model training would require data for all possible combinations of abnormal events to diagnose multi-abnormal events reliably. However,

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

https://doi.org/10.1016/j.net.2023.10.033
Received 26 May 2023; Received in revised form 27 September 2023; Accepted 22 October 2023
Available online 23 October 2023

the total number of combinations of abnormal events increases exponentially every time a label for one event is added, making it impractical to simulate a sufficient number of scenarios to produce the required amount of training data. In addition, as the amount of data increases, the burden on the model training also increases. The purpose of this study is to solve these problems by training a diagnosis model with a dataset consisting of only single abnormal events to have high diagnostic performance for untrained scenarios targeting multi-abnormal events, thereby removing the need to train multi-abnormal events.

This paper proposes an approach to achieve the diagnosis of two abnormal events occurring simultaneously by considering each event independently. First, feature selection is conducted applying an extra trees (or extremely randomized trees) classifier to select key parameter sets that are needed to diagnose each target abnormal event. Next, for multi-label classification, sub-models are trained by preprocessing all training data to include only the key parameter sets selected for specific target abnormalities. The trained sub-model set conducts binary classification for the occurrence of each target event. With this approach, we find an improved diagnostic accuracy of cases with multi-abnormal events. Based on the diagnosis improvements from the proposed approach, we expect that future operator support systems can cope with more various situations to secure stable NPPs.

The rest of this paper is organized as follows. Section 2 introduces the related background including recent trends in diagnosis model research and the need for multi-abnormal event diagnosis models. Section 3 describes the multi-abnormal event diagnosis framework using single abnormal event data, Section 4 introduces the case study of our work, and Section 5 presents the case study results showing high diagnostic performance from our approach as well as optimization steps. Sections 6 and 7 respectively discuss the results and summarize the conclusions of this study.

2. Background

This section introduces earlier studies on NPP state diagnosis applied with deep neural networks. In addition, the general approach to multi-label classification and its application to multi-abnormal event diagnosis, the goal of this study, are discussed.

2.1. Related work for NPP state diagnosis

Many models that predict NPP states have already been studied for the development of a system that can support operator tasks. Among them, those using data-based approaches such as artificial intelligence (AI) can operate efficiently by recognizing the state patterns resulting from problems that occur in NPPs with complex systems [3]. Embrechts et al. applied several neural network-based models, including a back-propagation algorithm, to identify malfunctions in NPPs [4]. Mo et al. diagnosed the type and severity of transients by applying a dynamic neural network aggregation model [5]. Santosh et al. developed an artificial neural network-based methodology to diagnose transients through data with reactor process parameters [6]. Bae et al. applied a multi-input multi-output strategy and long short-term memory networks to predict the trends of NPP parameters according to device controls by operators in an emergency situation [7]. Among the predictions of various NPP states, abnormal event diagnosis requires a detailed consideration of the changes in NPP parameters or specific malfunctions in components, so detailed classification levels are required for systems that support this task. Recently, a number of studies have been conducted to improve the performance of models dealing with the diagnosis of abnormal events. Kim et al. performed abnormal event diagnosis using principal component analysis and a gated recurrent unit model [8]. This algorithm achieved high performance by conducting the diagnosis through two stages corresponding to the procedure level and the detailed cause of the event. Kim et al. proposed a consistency check algorithm based on preprocessing using principal component analysis and an AI model using recurrent neural networks to diagnose abnormal events. The proposed algorithm diagnosed 34 abnormal events with an accuracy of about 0.98, and subsequently re-diagnosed 19 erroneous diagnosis cases correctly [9]. Yu et al. diagnosed sensor faults that occurred in various initial conditions based on preprocessing using principal component analysis and conditional flowcharts using random forest models [10]. In addition, the proposed algorithm enabled continuous learning and diagnosis. However, these previously studied data-driven models were trained with datasets limited to single abnormal events, and most of them were tested with previously trained scenarios. Considering this point, abnormal event diagnosis models using AI need to be improved before application to operator support systems in actual NPPs. This paper targets an approach to improve diagnosis models by achieving high performance in covering multi-abnormal events with no additional training.

2.2. Multi-label classification and multi-abnormal event diagnosis

In multi-label classification, each data instance can be correlated with multiple labels. Converting such multi-label data into single-label data in order to classify it into multiple labels can be accomplished in two ways [11,12]. In the first method, called label powerset, the multiple labels in the training dataset are converted into new single labels. In the other method, binary relevance, a given dataset with multiple labels is transformed to a dataset with multiple single labels for performing binary classification of each label. Certain binary classifiers derive relevance for the labels of particular data instances.

The current study applied multi-label classification to a dataset with NPP abnormalities. As systems or components within an NPP interact with each other, different abnormal events can affect the same systems or components; consequently, considering the dependencies of each abnormal event in a diagnosis model is very complex and requires careful consideration. In order to minimize the impact of these abnormality-specific dependencies on model diagnosis, we applied a transformation to independently handle each abnormal event included in the multi-abnormality dataset, just like two methods for multi-label classification mentioned above. When there are N kinds of single abnormal events, $N*(N–1)/2$ combined multi-abnormal events are derived even when the occurrence of only two kinds of abnormal events is considered. Therefore, it is costly to consider all multi-abnormal events that can occur by applying the label powerset method in the model training step, as this would require single labels for all possible multi-abnormal events. Accordingly, for multi-abnormal event diagnosis, it is advantageous to apply binary relevance to train existing single abnormality labels using a dataset of only single abnormal events.

3. Multi-abnormal event diagnosis framework

This section introduces our approach in this study. In order to diagnose multi-abnormal events that have occurred in an NPP, the following techniques can be used. First, feature selection can be applied to given data to select only the information necessary for the classification. Second, a model structure composed of multiple sub-models can be used to solve multi-label classification problems by working with binary relevance. In addition, the data used for model evaluation should have maximum diversity within the prediction range for multi-abnormal event diagnosis. Implementation details of the multi-abnormality datasets configuration, classifier for feature selection, and sub-models for diagnosis are given in Section 4.

3.1. Proposed approach with single-event training data

When an emergency situation occurs in an NPP, the plant operating parameters exceed the operating setpoints of the reactor protection system or engineering safety features, triggering the safety systems and numerous parameter changes. Differing from such emergency situations, abnormal events generally involve a specific component or system, and these events can typically be alleviated by operator actions without major changes to the NPP such as reactor trip or safety injections [13]. As abnormal events mostly induce changes in limited parameters related with the specific component or system, each individual event included in a multi-abnormal event affecting the components can have its own characteristic parameter changes. Examples of this are shown in Fig. 1. Based on this point, we try to diagnose abnormal events by attending to the more characteristic effects through the following algorithm. Our approach selects parameters using the classification importance of each feature calculated in the training of the machine learning algorithm to be able to detect the characteristic of each abnormal event. At this point, we label the target abnormality data as 1 and other abnormality data as 0 beforehand to train machine learning models for selecting features and to train sub-models for diagnosing each target abnormal event. Through this, feature selection is performed to include only the key features that characterize the target abnormal event. All feature-selected data are normalized within the maximum and minimum range of each parameter, and these preprocessed data are then used to train each sub-model for each target abnormal event. This training process is performed iteratively for each single abnormal event. For model diagnosis, key feature selection is performed iteratively for each abnormal event in the test data in the same way as in the training process. The transformed test data is diagnosed independently by sub-models for each single abnormal event. In other words, the diagnosis result by the proposed algorithm is derived in the form of 0 or 1 by all sub-models, that is, binary relevance, indicating whether each abnormal event has occurred. The entire model training and evaluation framework of the proposed approach is shown in Fig. 2.

3.1.1. Feature selection with extra trees classifier

Feature selection can improve the classification performance of a given model by selecting the features that are correlative and informative. In addition, it is expected that model simplification or model acceleration in classification tasks can be achieved by reducing the size of the data via feature selection. For this, the following methods can be applied: a filter method that uses correlation or statistical techniques between features and targets, a wrapper method that finds the optimized subset of features leading to the best performance of the machine learning model, and embedded methods that use an algorithm with a feature selection method already built in. The parameter changes according to one abnormal event occurring in an NPP are nonlinear and can be influenced by the automatic operation of the given component having a high malfunction degree. Therefore, the current approach suggests the feature selection method using the machine learning model because calculation using a single statistical technique is difficult considering the relationship between all abnormal events. Among

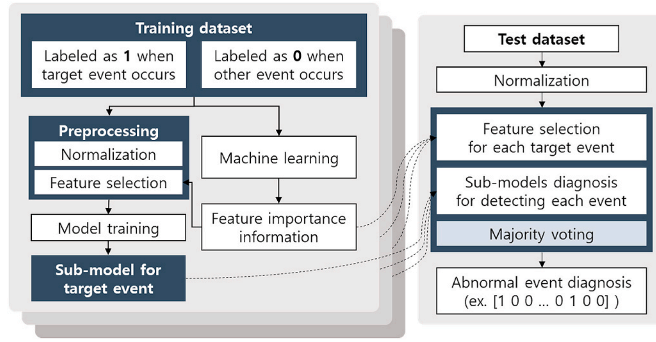

Fig. 2. Proposed approach to multi-abnormal event diagnosis.

model-based feature selection methods, one type of tree algorithm is called the extremely randomized trees classifier (extra trees classifier) [14]. For the machine learning algorithm, an extra trees classifier was selected, which has a more advanced architecture compared to the random forest classifier [15,16], a basic tree algorithm. The structure of this classifier is shown in Fig. 3.

Decision trees in the extra trees classifier are constructed from original training samples. Then in the test node, each tree gives a random sample of k features from the feature set, where each decision tree has to select the best feature to split the data according to mathematical criteria. This random feature sample produces multiple decorrelated decision trees. The model uses an ensemble learning technique that outputs classification results by aggregating the results of these decision trees. For each feature in this forest structure, the normalized feature importance is calculated based on mathematical

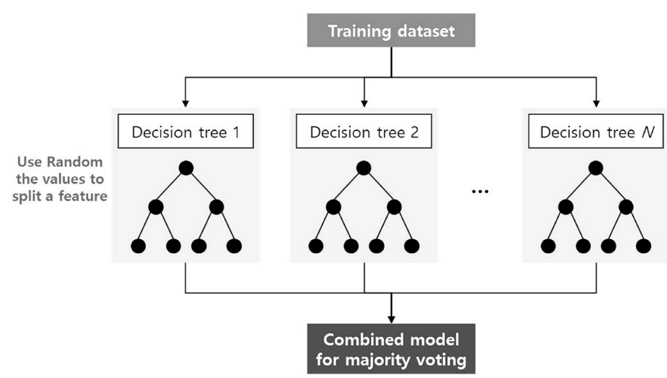

Fig. 3. Structure of an extra trees classifier.

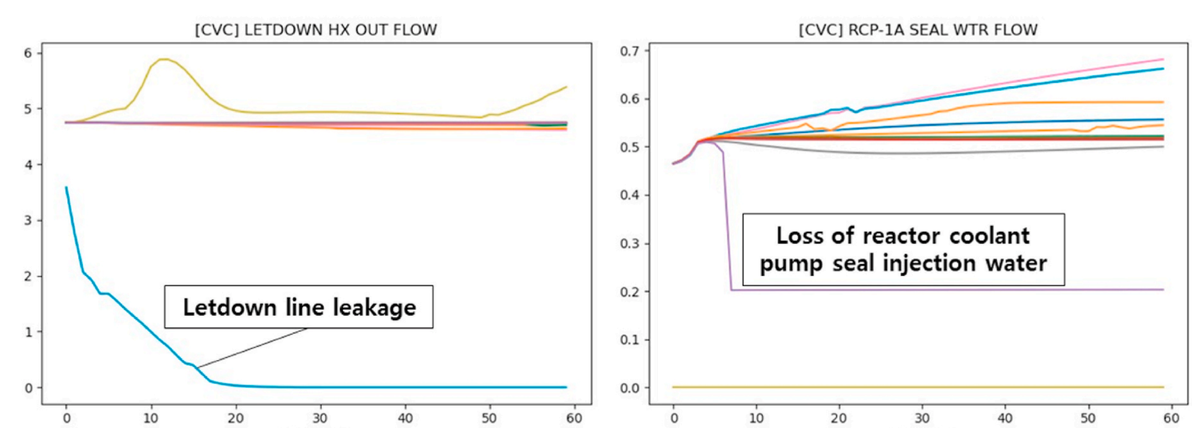

Fig. 1. Examples of characteristic parameter changes by an abnormal event.

criteria used to determine the feature splitting. Finally, feature selection is performed using this feature importance.

3.1.2. Two-Channel convolutional neural networks for NPP abnormality diagnosis

In an earlier study, Lee et al. applied an improved input structure of a basic convolutional neural network (CNN) to NPP abnormal event diagnosis [17]. The improved design consisted of two channels: the first input channel has parameter information at the current state, and the second input channel has parameter information with the changes in the parameters from 5 s ago. As a result, the trained model showed about 2% higher classification accuracy than a basic one for 10 abnormal events. This model structure makes it possible to reflect parameter change information over time due to abnormal events, which are the targets to be classified, and secure higher performance by considering time-series data of NPPs.

A typical convolution layer computes the convolution of the input data using filters of small size, such as (2*2) or (3*3). This convolution has advantages for input data in a two-dimensional format as it allows the model to train positional information for each pixel in the input data. However, raw data has a tabular form and does not have a two-dimensional form in a single time step. In an earlier study for the reference model, zero-padding was added to the current information and the parameter change information of sequentially listed parameters to convert them into a square shape for the two channels to solve this problem.

However, another means to solve this problem effectively is to apply a one-dimensional convolutional layer that can maintain the tabular characteristics of the raw data instead of making it into a square shape [18,19]. Fig. 4 shows the shape of the input data for this model. In this way, when applying a one-dimensional CNN to the classification of tabular data, higher performance can be expected by reflecting the characteristics of convolution calculation.

3.2. NPP event combination for multi-Labeling validation

For a deep learning model to achieve high performance, it has to be trained with sufficient and diverse data. But in the current case, the number of abnormal events that have occurred in actual NPPs is insufficient to make a training dataset for the model, and similarly, it is difficult to collect various cases for each abnormal event. In this regard, NPP simulators can be used to simulate desired abnormal events by injecting specific malfunction commands into various components. In addition, the intensity of the abnormal event can be set, such as the opening size of a valve or the leakage size of a tube due to a failure. As such, NPP simulators offer advantages in obtaining desired data, and accordingly can be used to collect single and multi-abnormal event data considering the lack of real-world data on multi-abnormal events.

The parameters affected by different abnormal events may be the same or different. If two different abnormal events occur at the same time but the intensity of one malfunction is relatively greater than the other, operator diagnosis may be concentrated on the abnormal event causing the relatively large change in the monitoring parameter, potentially leaving the other abnormal event unnoticed. In addition, if two abnormal events affecting the same monitoring parameters occur simultaneously, the parameters corresponding to different entry conditions may present mixed fluctuations, which could cause confusion in the operator diagnosis. Even in such situations, a diagnosis model has to prove capable of providing accurate diagnosis information. In other words, a high level of model diagnostic performance must be maintained in the case of abnormal events that are difficult for operators to diagnose. In order to consider numerous scenarios of concern, such as the above examples, multi-abnormal event scenarios can be conceived as shown in Fig. 5.

First, combinations of all abnormal events have to be specified taking into account the occurrence of two events. Next, it is required to make pairs of combinations of the intensity or fraction of each abnormal event. Multi-labeling validation of NPP abnormal events should be carried out by simulating such foreseeable scenarios.

4. Case study

The approach in this study seeks to diagnose multi-abnormal events by classifying whether each abnormal event has occurred or not in a test scenario. The multi-abnormal event diagnosis approach is designed to detect each event occurrence independently from parameter sets selected by the classifier. In this section, we introduce the case study to showing multi-abnormal event diagnosis with our approach.

4.1. Dataset configuration

As shown in Fig. 6, the 3KEYMASTER NPP simulator used for data collection follows the system structure of a general two-loop pressurized water reactor capable of producing 1400 MWe power [20]. When we constructed the ideal dataset to train and test the classification model, we considered a diversity of possible abnormal events through systematic scenarios.

Abnormal events were selected as malfunctions that can occur in each of three or more components in the primary system, secondary system, and auxiliary system. Table 1 lists the 15 abnormal events with corresponding labels selected for the data collection in this study.

In terms of the operating range in a scenario in which two abnormal events occur simultaneously, the range of malfunction intensity for each abnormal event is based on the minimum intensity at which an alarm or a symptom is sufficiently occurred and the maximum intensity at which an emergency situation does not occur due to a reactor trip within about twice the output time, as shown Fig. 7. For producing datasets, each

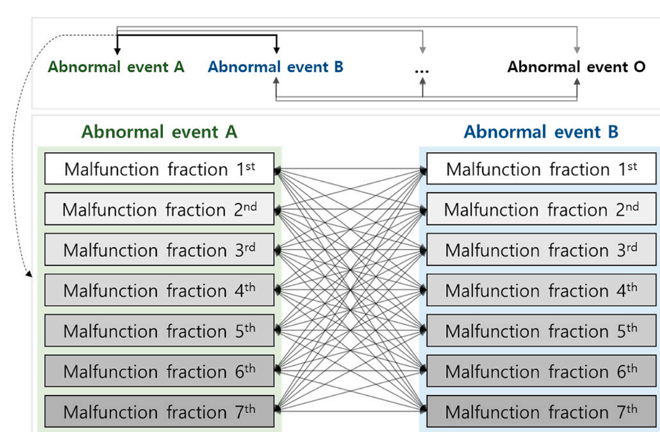

Fig. 5. Multiple scenarios in a multi-abnormality case with two abnormal events.

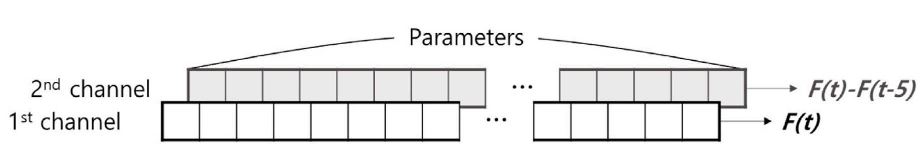

Fig. 4. Input structure of the two-channel CNN.

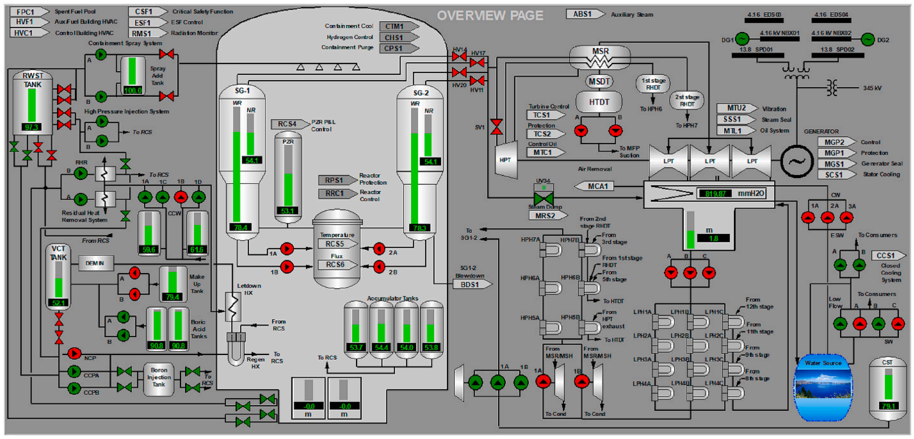

Fig. 6. 3KEYMASTER NPP simulator overview.

Table 1
Abnormal events information for data production.

Label	Abnormal event	Malfunction command	Min. fraction	Max. fraction
SGTL	Steam generator A tube leakage	IMF mf_RCS01	4	10
CHRG	Charging line break upstream of FT-121	IMF mf_CVC01	10	100
LTDN	Letdown line leakage inside containment	IMF mf_CVC05	100	1000
CDS	Loss of condenser vacuum	IMF mf_CON01	45	50
POSRV	POSRV[a] valve (HV456A) leakage	ICM fbv_PCV456A t:1	0.2	1
CWS	Circulating water tube leak in LP[a] condenser	IMF mf_CWS03	65	100
MSIV	Main steam isolation valve (HV14) valve positioner failure	ICM vmodABHV0014 t:4	0	0.3
RCP	Loss of reactor coolant pump seal injection water by valve (HV8351A) positioner failure	ICM movBBHV8351A t:4	0	0.03
MSS	Main steam header steam leakage	IMF mf_MRS09	2	3
PZR	Pressurizer spray valve (PV455B) positioner failure	ICM cmPos_PZRSpray1 t:4	70	100
CCW	Component cooling water service loop header leakage	IMF mf_CCW01	10	100
LFH	LP Feedwater heater 1A tube break	IMF mf_CON02	10	100
HFH	HP[a] Feedwater heater 5A tube break	IMF mf_MFW04	55	90
MFW	Main feedwater pump recirculation valve (FV1B) positioner failure	ICM vmodAEFV0001B t:4	0.45	0.7
TB	HP turbine control valve (CV1) positioner failure	ICM vmodACFCV0047 t:4	0	0.25

[a] POSRV, pilot-operated safety relief valve; LP, low pressure; HP, high pressure.

single abnormal event included 49 scenarios with malfunction intensities evenly divided across the criterion range for the event. This process was repeated twice to produce a training dataset and a test dataset. That is, each single-abnormality dataset was produced with 735 scenarios for model training and 735 scenarios for model testing. The NPP simulator was sampled every second during 60 time steps for 751

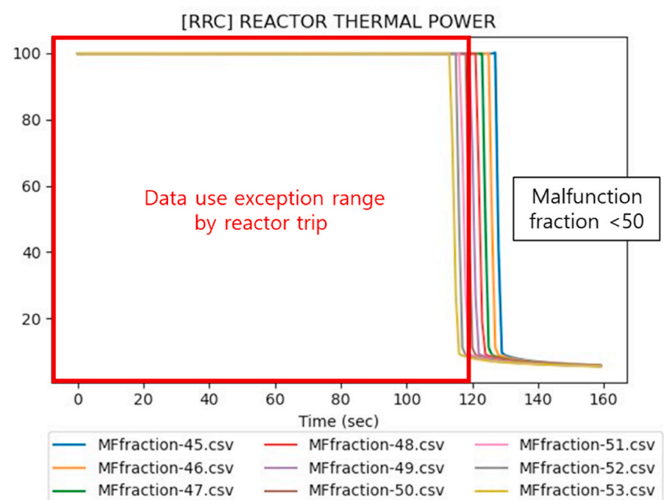

Fig. 7. Example of intensity setting with loss of condenser vacuum scenarios.

monitoring parameters that show changes for any of the 15 abnormal event cases.

This study conducted model training with only single-abnormality datasets, meaning that all of the multi-abnormality datasets produced were used for model evaluation. Two overall test datasets were considered. Test datasets 1 focused on the multi-abnormalities, where we considered a diversity of intensities, not only different kinds of malfunctions. A total of 105 multi-abnormal event cases were considered, which are all combinations of 2 of the 15 selected abnormal events occurring simultaneously. Next, 49 scenarios were created for each multi-abnormal event case considering the combinations of each possible malfunction range of the abnormal events. That is, multi-abnormality datasets were produced with a total of 5,145 scenarios for model evaluation. With test datasets 2, our purpose was to verify the diagnostic performance for cases in which there is a time difference between the occurrence of the two abnormal events. For this, we used 210 scenarios by considering the event order in the 105 combinations. The intensity of the events was fixed at the median value of the malfunction fraction, and the second event occurred 30 s after the first event.

4.2. Model structure for feature selecting and training

In the proposed approach, we first trained extra trees classifiers for each target abnormal event for feature selection. For this, each scenario was converted into 60 s of data corresponding to the whole time, giving a total of 44,100 data used for model training. To determine the number of selected features, we selected features above a certain ranking of feature importance. Other methods included selecting features above the average feature importance and selecting features using cross-validation during model training. The common detailed structure of each classifier for the 15 events is as follows.

- Number of trees: 200
- Criterion to measure the split quality: Gini impurity
- Maximum depth of the tree: Depth till all leaves contain less than 2 samples
- Number of features for the best split: square root of the total number

This paper targets NPP abnormal event diagnosis model training with time-series data. For this, a two-channel CNN that can reflect parameter changes over time in the NPP data (see Section 3.2) was used as a base model. As the input data needs a time step of 5 s to be formatted as two channels, each scenario taking 50 s is transformed into 55 data to be formatted. Thus, a total of 40,425 data is used for model training, of which 30% is splatted and used for validation at every epoch step.

The purpose of the sub-models is to basically perform a binary classification of whether the target abnormal event occurs or not. The models have a neural network architecture with only one convolutional layer and one fully connected layer to perform the simple classification. Also, since classification is performed on untrained data, which is multi-abnormality data, it is important to avoid overfitting the training data, which is single abnormal event data. For this, model training was terminated early if the log-loss of the validation dataset did not decrease for 10 epoch steps. In general models for multi-label classification, a sigmoid that has the property of returning a probability value for each label and that can represent two or more labels is typically used as the activation function, and binary cross-entropy is used as the loss function for the fully connected layer [21]. But here, since each sub-model in the proposed approach is tasked with the prediction of only a single label, softmax was used as the activation function and categorical cross-entropy was used as the loss function. The hyperparameters and structures for the sub-models are as follows.

- Epoch size: 50 epochs with early stopping
- Number of convolution layers: 1
- Filter size of the convolution layer: 3
- Filter number of the convolution layer: 32
- Activation function of the convolution layer: ReLU [22].
- Activation function of the fully connected layer: Softmax
- Loss function of the fully connected layer: Categorical cross-entropy
- Optimizer: Adam [23].

5. Test results

This section presents the diagnostic performance for multi-abnormality datasets of models trained with a dataset containing only single abnormal events. First, we show that the proposed approach can secure high diagnostic performance for multi-abnormal events compared to a base model that has no structural changes. Next, we verify the model and the feature selection technique used in the proposed approach for abnormal event diagnosis. We then conduct sensitivity studies adjusting the number of selected features to find the highest diagnostic accuracy for multi-abnormal events. Among them, Table 2 shows examples of the parameters for target events that were selected with only the minimum number by the extra trees classifier with recursive feature elimination with cross validation (RFECV). These selected parameter sets are related to the system in which each event occurred.

Table 2
Examples of the selected parameters for two events.

Label	LTDN	RCP
Selected parameter	Letdown HX[a] outlet flow	RCP[a]-2B seal water flow
	Letdown HX letdown outlet pressure	RCP-2A seal water flow
	Letdown HX outlet control valve (CV[a])	RCP-1B seal water flow
	Letdown HX outlet control valve (PRO[a])	RCP-1A seal water flow
	Regenerative HX charging outlet temperature	RCP-2A #1 seal leak flow
	Letdown HX letdown outlet temperature	RCP-1A #1 seal leak flow
	Letdown HX outlet to CCW[a] (CV)	RCP total seal injection flow(1)
	Letdown HX outlet to CCW[a] (PRO)	RCP total seal injection flow(2)
	Valve (PV131) in CVCS[a]	HI[a] pressurizer pressure channel-C
	Valve (TV130) in CVCS	Valve (HV8351A) in RCS[a]

[a] HX, heat exchanger; RCP, reactor coolant pump; CV, control variable; PRO, process variable; CCW, component cooling water; HI, high; CVCS, chemical and volume control system; RCS, reactor coolant system.

5.1. Improvement with feature selection and multi-label classification

The prior-studied two-channel CNN showed high diagnostic performance for a single abnormal event. We applied the proposed approach to this model and improved the diagnostic performance for multi-abnormal events. Below is the result with test datasets 1.

5.1.1. Multi-abnormal event diagnosis performance

One of the processes in the proposed approach is feature selection by an extra trees classifier, which is performed for each abnormal event. Another process is for each sub-model to independently classify whether each abnormal event has occurred. In order to confirm the performance improvement of multi-abnormal event diagnosis by the proposed approach, the diagnostic accuracy of each abnormal event occurrence type by each stage process was confirmed. The base model has the same structure as in Section 4.2, but here the activation function is changed to sigmoid for the distribution of prediction. The model marked ML has a structure composed of sub-models for multi-label classification without preprocessing (feature selection). The last model as the primary approach of this study, FS-ML, has a structure composed of sub-models for multi-label classification along with feature selection additionally performed with a set of selected parameters for each event.

As shown in Fig. 8, the model that independently diagnoses each abnormal event improves the accuracy by about 18% for the multi-abnormality dataset. The overall model, which added a process of selecting 50 parameters to diagnose the occurrence of each abnormal

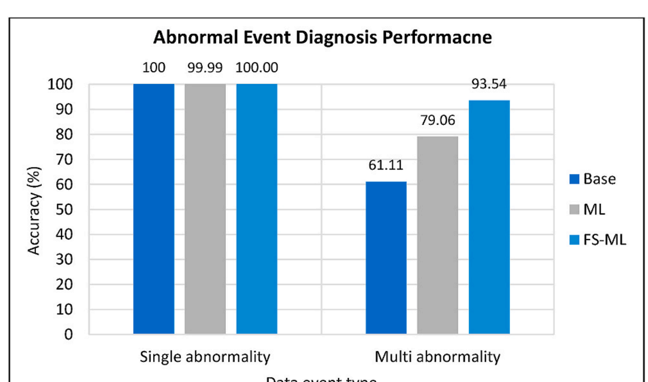

Fig. 8. Abnormal event diagnosis performance with the proposed approach.

event, improved the accuracy for the multi-abnormality dataset by 32.4% compared to the base model.

5.1.2. Diagnostic accuracy for each abnormal event

Although the model with the proposed approach was trained on single-abnormality datasets, it showed 93.54% accuracy on the multi-abnormality dataset. The results for the multi-abnormality data including each abnormal event are shown in Table 3. It used for accuracy as the rate was measured when the model diagnosed both abnormal events included in the multi-abnormality data.

As shown in the previous section, the performance for each single abnormal event was mostly maintained in the overall model as well as in the base model. As a result of the evaluation of multi-abnormality data, the diagnostic performance of most cases was improved in the model that applied only the independent diagnosis process excluding feature selection, but the accuracy of the CHRG cases was rather low, by 2.16%. On the other hand, looking at the results applying the overall model, it was confirmed that the multi-abnormal event diagnosis performance for all abnormal conditions was improved. In particular, the accuracy improved by 57.23% for the multi-abnormality dataset for PZR events and by 74.83% for that for RCP events, as marked in bold in Table 3. Finally, it was confirmed that the model achieved a minimum diagnostic accuracy of 82.62% and a maximum of 99.84% for each abnormal event.

5.2. Sensitivity studies

In this study, we tried to improve the two-channel CNN, which is a verified abnormal state diagnosis model, and to confirm this, neural networks were selected as sub-models. In addition, for feature selection, the necessary parameters were selected through an extra trees classifier with improved performance compared to the random forest classifier. In this section, each selected technique was verified to have an appropriate level of results compared to other techniques. Additionally, we tried to improve the overall diagnostic accuracy by changing the number of parameters selected for each sub-model.

5.2.1. Comparison with other techniques

Sub-models should be able to diagnose not only the occurrence of target abnormal events in single-abnormality datasets but also the occurrence of target abnormal events in datasets containing parameter information reflecting multi-abnormal events. In order to confirm the performance of the two-channel CNN for this purpose, other representative binary classifiers or simple neural networks were selected and evaluated with test datasets 1. Among the selected models, the support vector machine [24] is a basic model with a poly kernel, and the one-channel convolutional neural network has the same hyperparameters as the two-channel CNN but without the second channel reflecting parameter change information over time. In the case of the simple neural network, the number of nodes in the dense layer is adjusted so that the total number of parameters is similar to that of the base model. Table 4 shows a comparison of the results of each diagnosis model. In addition, by comparing several classifiers for the feature selection technique, we confirmed whether the extra trees classifier selected in this study has superior performance. Since the purpose of these classifiers is to obtain information on feature selection rather than to diagnose the event directly, machine learning models were selected as a comparison group. A linear support vector machine and logistic regression, which are binary classification models, a random forest classifier, which is a lower-level tree model compared to the extra trees classifier, and LightGBM [25], which showed the highest results in NPP event diagnosis in an earlier study [26], were tested with test datasets 1. The selected parameters were composed of only parameters whose feature importance by each classifier was higher than the mean value. Table 5 shows the comparison results.

The sub-models showed lower diagnostic accuracies, under 87%, for multi-abnormality data across all comparative models considering inputs with single-time-point parameter information. In the case of the one-channel CNN, the single-abnormality diagnosis accuracy was maintained at about 99%, but the multi-abnormality diagnosis accuracy was lower at 77.20%. In contrast, the two-channel CNN selected in this study achieved a relatively high diagnostic performance of 93.54%.

5.2.2. Feature number and time difference verification

In this work, we tried to secure diagnostic performance for multi-abnormal events by selecting appropriate parameters for each sub-model to provide data with limited parameter information. On this point, we can further improve the model performance by confirming the optimal amount of parameter information to include. In the previous

Table 4
Diagnosis performance with the extra trees classifier.

Classification model type	Accuracy for single abnormalities (%)	Accuracy for multiple abnormalities (%)
Support vector machine	98.12	87.97
Artificial neural network	98.75	84.23
One-channel convolutional neural network	98.74	77.20
Two-channel convolutional neural network	100.00	93.54

Table 3
Abnormal event diagnosis performance comparison.

Model	Base		ML			FS-ML		
Model type	Single model		Sub-model set			Sub-model set		
Feature selection	None		None			Higher than mean importance		
Event type	Single	Multi	Single	Multi	Imp. of multi	Single	Multi	Imp. of multi
SGTL	100.00	41.29	100.00	52.95	11.66	100.00	82.62	41.33
CHRG	100.00	81.58	100.00	79.42	−2.16	100.00	96.75	15.17
LTDN	100.00	76.57	100.00	82.68	6.11	100.00	93.98	17.42
CDS	100.00	69.27	100.00	88.21	18.94	100.00	99.77	30.50
POSRV	100.00	68.95	99.93	82.65	13.71	100.00	92.77	23.83
CWS	100.00	76.30	100.00	89.50	13.20	100.00	98.91	22.61
MSIV	100.00	76.26	100.00	91.88	15.62	100.00	97.70	21.44
RCP	100.00	25.00	100.00	86.93	61.92	100.00	99.84	**74.83**
MSS	100.00	64.53	100.00	76.29	11.76	100.00	83.97	19.44
PZR	100.00	26.85	100.00	47.38	20.53	100.00	84.08	57.23
CCW	100.00	75.98	99.96	89.57	13.58	99.93	97.77	21.78
LFH	100.00	46.53	100.00	78.21	31.68	100.00	95.93	49.40
HFH	100.00	67.65	100.00	84.28	16.63	100.00	94.31	26.66
MFW	100.00	70.27	100.00	91.19	20.92	100.00	94.97	24.70
TB	100.00	49.62	100.00	64.72	15.10	100.00	89.68	40.06

Table 5
Diagnostic performance of the two-channel CNN.

Feature selection method	Accuracy for single abnormalities (%)	Accuracy for multiple abnormalities (%)
Linear support vector machine	100.00	90.80
Logistic regression	100.00	92.03
Random forest classifier	100.00	92.21
Extra trees classifier	100.00	93.54
LightGBM	100.00	89.80

section, selection was made based on parameters having a feature importance higher than the criterion, or mean value.

In this section, diagnostic performance was checked using test datasets 1 for different numbers of parameters for each sub-model, namely 50, 100, and 150 selected in order of highest feature importance, including the previous criterion. In addition, we tried to automatically select the number of parameters required for each abnormal event. With RFECV, while training each classifier, the number of parameters with the best performance is determined while removing the least important parameters one by one. Fig. 9 shows the accuracy of the model for diagnosing abnormal events based on these selected parameters. From the above results, it was confirmed that when 100 parameters for each sub-model were selected, the diagnostic accuracy for multi-abnormal events was the highest at 94.21%. Moreover, we found that a high accuracy of 91% or more was achieved in all cases of feature selection.

Table 6 shows the accuracy for the data containing each abnormal event including multi-abnormality cases. The model using 100 parameters per sub-model, which is the one with the highest diagnostic performance, maintained about 100% performance for each single abnormal event. In addition, for each abnormal event, a performance of about 85% or more was secured for multi-abnormal events.

We then evaluated the model that showed the highest accuracy of 94.21% using test datasets 2. Test datasets 2 constituted 5,250 data corresponding to multi-abnormal situations with different event timings, namely a second malfunction injection 30 s after the first malfunction. Fig. 10 shows the results for multi-abnormal event diagnosis comparing the two test datasets. The model diagnosed test datasets 2 with an accuracy of 94.97%, which is similar to the result for test datasets 1, even though the events occurred at different times.

6. Discussion

In order to discuss the appropriateness of the proposed approach for use as an NPP abnormality diagnosis model, the following points about the experimental results should be mentioned. In Section 5.1, it was shown that each process in the proposed approach secured the ability to cope with diagnosis with 93.54% accuracy when a multi-abnormal event occurred. However, the frequency of single abnormality occurrence is higher than that of multi-abnormality occurrence in actual NPPs. In other words, it is important for the model to maintain the ability to accurately perform existing tasks while gaining the ability to cope with unexpected abnormal events such as multi-abnormal events. In this regard, the overall algorithm showed almost equivalent performance to the base model, close to 100% accuracy, for single abnormality diagnosis.

In Section 5.2.1, the model diagnostic performance for the target data was verified by comparing the experimental results of various classification models and the two-channel CNN selected for the proposed approach. Through this, it can be confirmed that the diagnostic performance for multiple abnormalities is high when considering inputs with not only parameter information at a single point in time but also parameter change information together. In addition, sensitivity studies confirmed that the use of the extra trees classifier for feature selection can lead to the highest diagnostic performance among the tested techniques by selecting appropriate parameters for multi-abnormality diagnosis. Further studies can be conducted for more detailed comparisons by applying more classification models or feature selection techniques to reach the best performance; here, representative techniques and model hyperparameters of reasonable levels were selected for a general comparison. Through this, it can be inferred that the techniques included in the proposed approach can already show sufficient performance in multi-abnormality diagnosis at the same or higher level than other techniques.

The performance of an artificial neural network is influenced by the information contained in the data to be learned. Therefore, it is necessary to limit confounding information in order to improve the performance of models that classify multi-abnormality data, which was considered in the sensitivity analysis of Section 5.2.2. Among the tested feature selection methods, RFECV represents a detailed means of automatically selecting parameters for each event to maximize the diagnostic performance for a single abnormal event. However, as the results showed, models using the mean value for the threshold or using RFECV did not achieve higher performance than the models that simply selected a fixed number of parameters. Feature selection as used in this study has

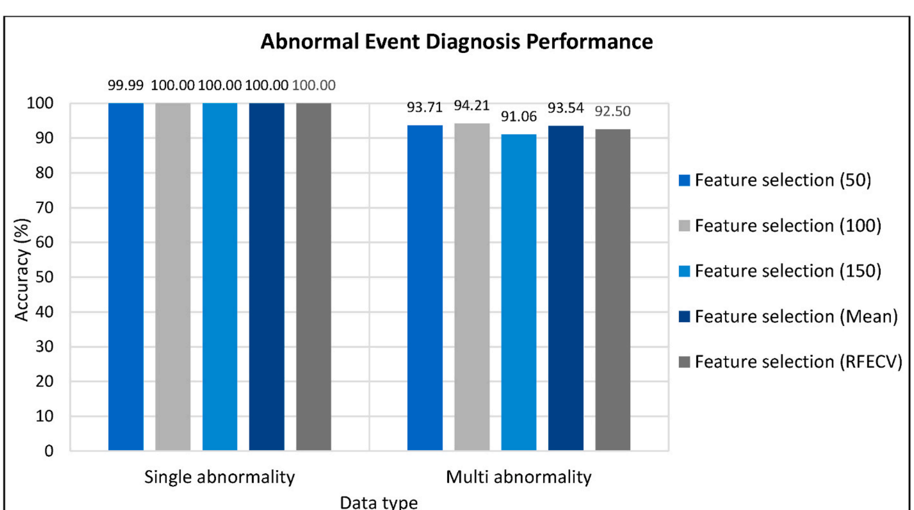

Fig. 9. Abnormal event diagnosis performance by number of selected features.

Table 6
Accuracy with test datasets 1 for each abnormal event by feature selection number.

Feature selection	Feature selection number									
Factor	50		100		150		>Mean		RFECV	
Event type	Single	Multi	Single	Multi	Single	Multi	Single	Multi	Single	Multi
SGTL	100.00	91.17	100.00	88.48	100.00	76.14	100.00	82.62	100.00	79.42
CHRG	100.00	99.86	100.00	97.82	100.00	95.06	100.00	96.75	100.00	97.81
LTDN	100.00	92.93	100.00	95.95	100.00	91.82	100.00	93.98	100.00	94.90
CDS	100.00	97.75	100.00	97.80	100.00	95.21	100.00	99.77	100.00	96.88
POSRV	100.00	93.12	100.00	93.28	100.00	90.75	100.00	92.77	100.00	92.01
CWS	100.00	99.98	100.00	99.37	100.00	98.25	100.00	98.91	100.00	99.58
MSIV	100.00	95.90	100.00	97.46	100.00	96.93	100.00	97.70	100.00	97.27
RCP	100.00	99.99	100.00	99.59	100.00	99.72	100.00	99.84	100.00	99.96
MSS	100.00	82.64	100.00	85.02	100.00	84.31	100.00	83.97	100.00	81.14
PZR	100.00	86.80	100.00	87.73	100.00	87.70	100.00	84.08	100.00	87.61
CCW	99.85	99.88	99.96	97.91	99.93	96.18	99.93	97.77	99.96	96.30
LFH	100.00	91.70	100.00	92.50	100.00	85.13	100.00	95.93	100.00	93.16
HFH	100.00	92.67	100.00	94.85	100.00	92.63	100.00	94.31	100.00	94.72
MFW	100.00	94.68	100.00	95.33	100.00	90.43	100.00	94.97	100.00	93.81
TB	100.00	86.62	100.00	90.06	100.00	85.56	100.00	89.68	100.00	82.89

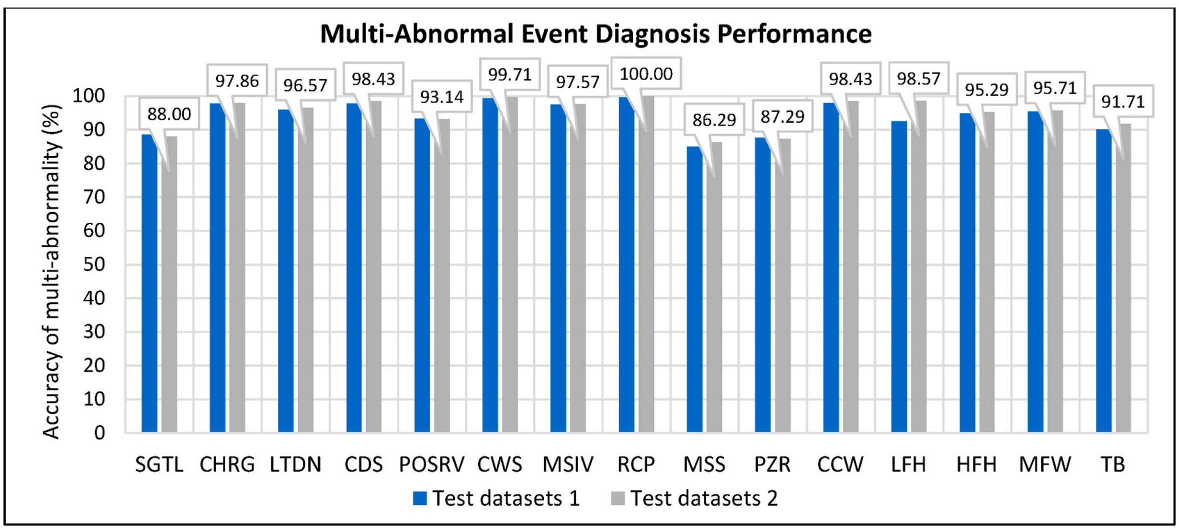

Fig. 10. Multi-abnormal event diagnosis performance comparison between each test datasets.

the purpose to improve the model performance not only for the classification of single abnormal events but also for the classification of multi-abnormal events that are not included in the training data. Therefore, when we apply feature selection using a model trained on a dataset containing only single abnormal events, automatic selection or standard selection using the average feature importance as a threshold may not be appropriate. In other words, they are not intended to be unconditionally suitable for the diagnosis of multi-abnormal events. In conclusion, while we may predict an appropriate number of parameters for each abnormality data, we cannot predict which parameters are appropriate for untrained data such as multi-abnormality data. However, through the analysis in Section 5.2.2, it can be inferred that information from about 100 selected parameters is helpful for diagnosing abnormality occurrence in each model. The results of test datasets 2, moreover, showed that the sub-models could diagnose each target event regardless of the time of occurrence of the two events.

7. Conclusion

Operators enter the appropriate abnormal operating procedure based on the plant information obtained following a single abnormal event, for which they need to recognize the plant condition changes. However, operators may find their diagnosis tasks to be relatively difficult in cases with changes caused by more complex events. Therefore, the artificial intelligence models used in systems developed to support operator diagnosis tasks for application in actual NPPs need to be able to cope with these situations. Despite this requirement, though, it remains difficult for the model training process to consider in advance all possible abnormal events at an NPP. In this study, we tried to diagnose multi-abnormal events in which two events occur simultaneously by training the model on only single abnormal events. For this, a multi-label classification model with sub-models using two-channel CNNs was proposed, and preprocessing was performed with 100 parameters selected through an extra trees classifier for each sub-model. The model following the proposed approach ultimately achieved a performance of 94.21% for multi-abnormality diagnosis while maintaining high performance for single abnormality diagnosis. This result represents a 33.1% improvement over the base model.

These improvements make it possible to diagnose most multi-abnormal events with only a small amount of training data. From a larger perspective, the possibility of diagnosis improvement for scenarios other than what are trained increases the reliability of the diagnosis support system, and thus the proposed approach can be expected to be applied to actual NPPs. However, additional studies are needed to overcome the discrepancy between real and simulator data before application to real NPPs. Prior research has shown that models trained with specific simulator data were able to diagnose other types of data or data added with noise [27,28]. Therefore, further work can be conducted to improve the structure of the sub-model in the proposed algorithms, in particular. The proposed approach showed high diagnostic performance for each abnormal system by controlling the quantity of information in the data and by varying the structure of the model. But since we do not fully control the feature selection by the tree classifier, we may not be able to completely rule out coupled effects in certain cases, such as abnormalities in matched systems. Therefore, if high accuracy for each multi-abnormal event cause is needed individually, additional training for such cases should be conducted in the future.

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.

Acknowledgements

This work was supported by National Research Foundation of Korea (NRF) grants funded by the Korean government (Ministry of Science and ICT) (RS-2022-00144042 and RS-2022-00144150).

References

[1] S.J. Lee, P.H. Seong, Development of an integrated decision support system to aid cognitive activities of operators, Nuclear Engineering and Technology 39 (2007) 703.
[2] A. Mathew, P. Amudha, S. Sivakumari, Deep learning techniques: an overview, advanced machine learning technologies and applications, Proceedings of AMLTA 2021 (2020) 599–608.
[3] J. Ma, J. Jiang, Applications of fault detection and diagnosis methods in nuclear power plants: a review, Progress in Nuclear Energy 53 (2011) 255–266.
[4] M.J. Embrechts, S. Benedek, Hybrid identification of nuclear power plant transients with artificial neural networks, IEEE Transactions on Industrial Electronics 51 (2004) 686–693.
[5] K. Mo, S.J. Lee, P.H. Seong, A dynamic neural network aggregation model for transient diagnosis in nuclear power plants, Progress in Nuclear Energy 49 (2007) 262–272.
[6] T. V Santosh, A. Srivastava, V.V.S.S. 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 (2009) 759–762.
[7] J. Bae, G. Kim, S.J. Lee, Real-time prediction of nuclear power plant parameter trends following operator actions, Expert Syst Appl 186 (2021), 115848.
[8] J.M. Kim, G. Lee, C. Lee, S.J. Lee, Abnormality diagnosis model for nuclear power plants using two-stage gated recurrent units, Nuclear Engineering and Technology 52 (2020) 2009–2016.
[9] G. Kim, J.M. Kim, J.H. Shin, S.J. Lee, Consistency check algorithm for validation and re-diagnosis to improve the accuracy of abnormality diagnosis in nuclear power plants, Nuclear Engineering and Technology 54 (2022) 3620–3630.
[10] Y. Yu, M. Peng, H. Wang, Z. Ma, S. Cheng, R. Xu, A continuous learning monitoring strategy for multi-condition of nuclear power plant, Ann Nucl Energy 164 (2021), 108544.
[11] R.B. Pereira, A. Plastino, B. Zadrozny, L.H.C. Merschmann, Categorizing feature selection methods for multi-label classification, Artif Intell Rev 49 (2018) 57–78.
[12] N. Spolaôr, E.A. Cherman, M.C. Monard, H.D. Lee, A comparison of multi-label feature selection methods using the problem transformation approach, Electron Notes Theor Comput Sci 292 (2013) 135–151.
[13] S. Kim, J. Park, S. Han, H. Kim, Development of extended speech act coding scheme to observe communication characteristics of human operators of nuclear power plants under abnormal conditions, J Loss Prev Process Ind 23 (2010) 539–548.
[14] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, Mach Learn 63 (2006) 3–42.
[15] T.K. Ho. Random decision forests, IEEE, 1995, pp. 278–282.
[16] L. Breiman, Random forests, Mach Learn 45 (2001) 5–32.
[17] G. Lee, S.J. Lee, C. Lee, A convolutional neural network model for abnormality diagnosis in a nuclear power plant, Appl Soft Comput 99 (2021), 106874.
[18] R. Shwartz-Ziv, A. Armon, Tabular data: deep learning is not all you need, Information Fusion 81 (2022) 84–90.
[19] M. Azizjon, A. Jumabek, W. Kim, 1D CNN based network intrusion detection with normalization on imbalanced data, in: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), IEEE, 2020, pp. 218–224.
[20] 3KEYMASTER Simulator, Western Service Corporation, Frederick, MD, USA, 2013.
[21] T. Durand, N. Mehrasa, G. Mori, Learning a deep convnet for multi-label classification with partial labels, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 647–657.
[22] V. Nair, G.E. Hinton, Rectified linear units improve restricted Boltzmann machines, in: Proceedings of the 27th International Conference on Machine Learning, ICML-10), 2010, pp. 807–814.
[23] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, 2014. ArXiv Preprint ArXiv:1412.6980.
[24] C. Cortes, V. Vapnik, Support-vector networks, Mach Learn 20 (1995) 273–297.
[25] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, Lightgbm: A highly efficient gradient boosting decision tree, Adv Neural Inf Process Syst 30 (2017).
[26] H.S. Ryua, K.N. Yub, Y.G. Kim, Development to Diagnose Model of Abnormal Status in Nuclear Power Plant Operation using Machine Learning Algorithms, in: Korean Nuclear Society Virtual Spring Meeting, July 9-10, 2020.
[27] J.H. Shin, J.M. Kim, S.J. Lee, Abnormal state diagnosis model tolerant to noise in plant data, Nuclear Engineering and Technology 53 (2021) 1181–1188.
[28] D. Lee, J. Kim, Concept of robust AI with meta-learning for accident diagnosis, in: Korean Nuclear Society Spring Meeting, Jeju, Korea, May 19-20, 2022.