Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio
©2023 ESREL2023 Organizers. *Published by* Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P597-cd

Multi-label Classification with Embedded Feature Selection for Complex Abnormal Event Diagnosis

Ji Hyeon Shin

Seung Jun Lee*

Department of Nuclear Engineering, Ulsan National Institute of Science and Technology, 50, UNIST-gil, Ulsan 44919, Republic of Korea. E-mail: sjlee420@unist.ac.kr

A nuclear power plant is the largest electrical power generation system composed of hundreds of components. When an abnormal situation occurs in a nuclear power plant, operators have to perform an appropriate diagnosis to alleviate the plant state. This abnormal event diagnosis process is based on the alarms and symptoms described in the abnormal operating procedures. However, when two or more abnormal events occur simultaneously, the plant parameters may show complex changes unlike the alarms and symptoms described. Abnormal event diagnosis models can be helpful greatly to operators when they can provide diagnostic information in more difficult situations such as these. In this study, the diagnostic performance of the existing artificial neural network model was improved by applying embedded feature selection to classify complex abnormal events. An embedded feature selection uses the feature importance of parameters used when a pre-prepared machine learning classifier trains a dataset. The parameters selected through this method only the characteristic parameters for each event so that the artificial neural network model can efficiently perform diagnosis. These results enable the abnormal state diagnosis model to provide diagnostic information to operators even in complex situations. In conclusions, this approach can increase the applicability of the diagnostic model using artificial neural networks to the actual operator support system for safer actual nuclear power plants.

Keywords: Nuclear power plant, abnormal event diagnosis, convolutional neural network, multi-label classification, embedded feature selection.

1. Introduction

Operators must recognize abnormal problems occurring in nuclear power plants (NPPs) to prevent risks leading to emergency accidents. These diagnosis tasks are performed by operators based on the numerous monitoring parameters and alarms that exist in NPPs. If more than one abnormal event occurs, each abnormal event may cause changes in several monitoring parameters, which may cause difficulty in diagnosis by operators. Therefore, the artificial intelligence model developed to support the diagnosis task of the operator must have the performance that can help even in these difficulties. Recently, many artificial intelligence models for multi-label classification have been studied. Using a model combined a convolutional neural network with long short-term memory or using a graph neural network to simultaneously predict multiple labels. However, for this case, training the artificial intelligence model is difficult because it requires a lot of physical cost to produce considering all the complex abnormal event scenarios.

This study proposed to improve diagnostic performance for complex abnormal event data by performing embedded feature selection using individual abnormal event data sets. For feature selection, the feature importance given in the Extremely randomized trees classifier (extra-trees classifier) was used. Each parameter selected data is diagnosed by sub-models that perform binary classification on target abnormal events.

2. Methodology

Data pre-processed through feature selection has a great influence on improving the performance of the model. Among feature selection methods, embedded feature selection is a method of selecting features by applying criteria appropriate to the feature importance trained by a machine learning model. Various machine learning models may be used, among them, extra trees classifier is an ensemble learning technique that uses multiple decision trees to output its classification result [1]. It is similar to the random forest classifier, but differs in the manner of construction of the decision trees in the forest. This class fits a number of randomized decision trees on various subsamples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. In addition, this classifier what we used in this study is advantageous for feature selection because Extra-trees random selection of a split value makes evaluate features from a wider perspective than another tree model.

3. Experimental Setup

In this study, we approach improving the diagnostic performance of each abnormal event through feature selection [2]. In addition, complex abnormal event diagnosis was performed by individually detecting each abnormal event as a target through multiple sub-models.

3.1. *Dataset*

All dataset used in this study were acquired through the NPP simulator made by Western Services Corporation [3]. The simulator is based on a generic pressurizer water reactor, and it can simulate abnormal events by injecting malfunctions into a desired component or system. The simulated abnormal events are introduced in Table 1. The training dataset was acquired with 25 scenarios composed of malfunction fractions at equal intervals for each abnormal event. One scenario involves acquiring data to include parameter information for 60 seconds for 797 parameters.

Table 1. Abnormal Event Description with Training Datasets.

Label Name	Event Name	Malfunction Description	Malfunction Fraction Range (Min, Max)
P2S2	CHRG	Charging line break	10, 100
P3S1	LTDN	Letdown line leak	100, 1000
P4S1	CDS	Loss of condenser vacuum	45, 50
P5S1	POSRV	Pilot operator safety relief valve leak	0.2, 1
P7S1	CWS	Circulating water tube leak	65, 100
P9S3	RCP	Reactor coolant pump seal injection water loss	0, 0.03
P11S1	PZR	Pressurizer spray valve positioner failure	70, 100
P12S1	CCW	Component cooling water service loop header leak	10, 100

- Test Dataset 1

Test data set 1 was created using 25 scenarios with a malfunction fraction different from the training data for each abnormal event. In addition, 25 scenarios were produced by applying different malfunction fractions to each case in which two abnormal events occurred simultaneously. That is, test data set 1 consists of a data set for a total of 200 common abnormal event scenarios and a data set for a total of 700 complex abnormal event scenarios.

- Test Dataset 2

Five test cases were created through scenarios in which two abnormal events with high systematic correlation occurred at different times. For example, the circulating water system deals with the circulation of seawater for cooling the condenser component. Alternatively, the water level of the pressurizer component is affected by the charging flow. Test dataset 2 was selected considering the correlation between these systems or components as domain knowledge.

The model was trained using only the single abnormal event data set and evaluated with each test data set. Through this, we can confirm that the proposed model can correctly diagnose complex abnormal event scenarios that can cause confusion in the operator diagnosis.

3.2. *Classification Model Structure*

Sub-models were trained with a binary classification system to diagnose the occurrence of each target abnormal event. Models for diagnosing the occurrence of each target event are trained with label '1' for the training data of the target event and label '0' for the training data of others. Assuming that the multi-abnormal event occurring are unknown, we cannot clarify the threshold of the cluster for the target event that may contain that case. Therefore, it is necessary to classify events by specifying what occurred and what did not occur through the supervised model with binary classification. They used a convolutional neural network (CNN) model with one convolution layer to perform simple classification. Table 2 shows the hyperparameters of the used model. Softmax for activation function and categorical cross-entropy were selected for binary classification of the model.

Table 2. Classification Model Hyperparameter.

Layer	Hyperparameter Description
Filter number of convolution layer	32
Kernel size of convolution layer	3
Activation function of convolution layer	ReLU [4]
Activation function of dense layer	Softmax
Loss function	Categorical cross-entropy
Optimizer	Adam [5]
Number of epochs	100
Additional function	Early stopping monitored validation loss with 10 patience

3.3. *Training Process*

Extra-trees classifier trained using a given dataset to detect a specific abnormal event can return a Gini importance for that event. Among them, the main parameters for the event are selected based on the mean Gini-importance [6]. In order to train a sub-model that detects the event, the training dataset is pre-processed by selecting key features for the target event and normalizing. The pre-processed training dataset is used to train whether the event occurs in sub-models. Through this process, eight sub-models are trained to detect each abnormal event. Figure 1 shows the sub-model training process. The test dataset is diagnosed by voting on the results of sub-models.

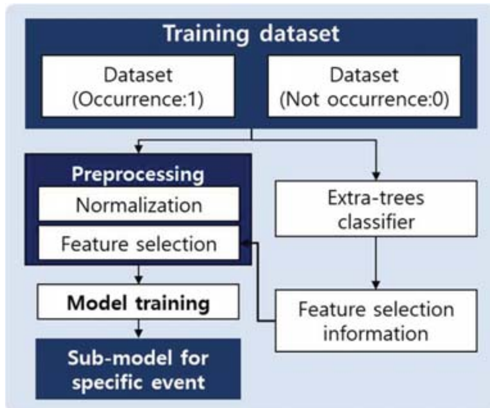

Fig. 1. Sub-model Training Process.

4. Results

In this section, we compared the proposed approach with a general CNN model with the same model structure. To enable diagnosing two events, we changed the activation function of the general model with 'sigmoid' and used 0.5 threshold at predicted values, for which is classical multi-label classification methods. The diagnostic results of test dataset 1 using the approach proposed in this study are shown in Table 3. We purposed to improve the performance of the model in diagnosing both abnormal events. Therefore, accuracy about two events scenarios was evaluated as Eq. (1) using both events rather than using each event. That is, we calculated true positive (TP) for the number of cases correctly diagnosed two combined events, true negative (TN) for the number of cases correctly diagnosed to the normal state, false positive (FP) for the number of cases diagnosed as one or more events despite the normal state, and false negative (FN) for the number of cases diagnosed as normal state despite events occurrence.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Table 3. Results about Test Dataset 1.

Model	Accuracy about One Event Scenario (%)	Accuracy about Two Events Scenario (%)
General CNN	99.38	69.23
Our approach	98.86	92.12

The approach proposed in this study has a single abnormal event diagnosis performance that is almost the same as that of the general model. In addition, the model improved diagnostic performance by about 22.89 % points when two events occurred simultaneously. Fig. 2 shows the improvement in diagnostic accuracy for each case in which the two events occur simultaneously when using the proposed approach than by using the general model.

	CHRG	LTDN	CDS	POSRV	CWS	RCP	PZR	CCW
CHRG		15.67	18.00	-38.87	5.53	11.33	90.20	61.73
LTDN	15.67		3.33	-0.73	0.07	26.13	92.00	5.53
CDS	18.00	3.33		20.20	41.87	-1.33	75.80	2.93
POSRV	-38.87	-0.73	20.20		4.07	41.87	58.80	2.93
CWS	5.53	0.07	41.87	4.07		0.13	20.53	1.27
RCP	11.33	26.13	-1.33	41.87	0.13		60.93	2.47
PZR	90.20	92.00	75.80	58.80	20.53	60.93		18.53
CCW	61.73	5.53	2.93	2.93	1.27	2.47	18.53	

Fig. 2. Accuracy improvement for each multi-abnormal event in the proposed model compared to the general model.

The proposed approach improved diagnostic accuracy in most cases. It even improved by up to 92% points for the PZR-LTDN case. However, it showed rather lower performance only for the CHRG-POSRV case.

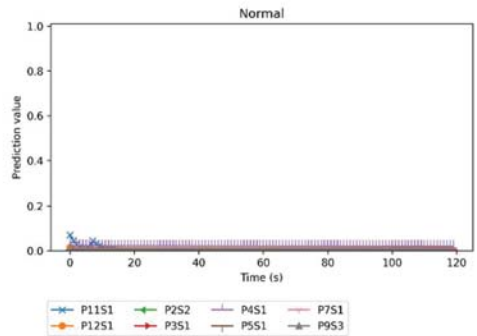

Fig. 3. Diagnosis Result of Normal State Scenario.

Although the proposed model did not train with data about the normal state, it did not classify the normal state into other events, as shown in Fig. 3. Fig. 4–8 below show the diagnostic results of the proposed approach on test dataset 2.

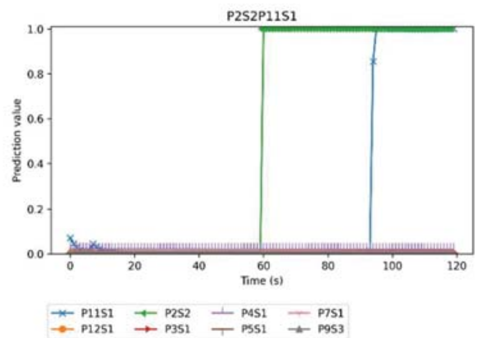

Fig. 4. Diagnosis Result of CHRG-PZR Event Scenario.

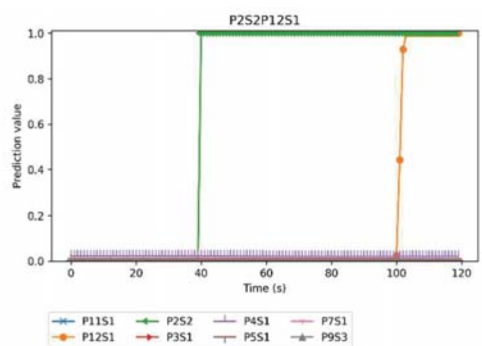

Fig. 5. Diagnosis Result of CHRG-CCW Event Scenario.

In the scenario of Fig. 4, the PZR event occurs 30 seconds later than the CHRG event. Compared to the existing model, which has a low accuracy of about 4.93% for the occurrence of CHRG-PZR events, the proposed approach shows that the two events were diagnosed well even though they occurred at different times. In the scenario of Fig. 5, the CCW event occurs 60 seconds later than the CHRG event. Also, for this scenario, the existing model has a low accuracy of 33.73%, but the proposed approach shows that it is well diagnosed.

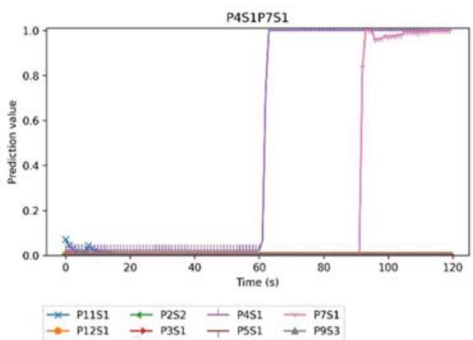

Fig. 6. Diagnosis Result of CDS-CWS Event Scenario.

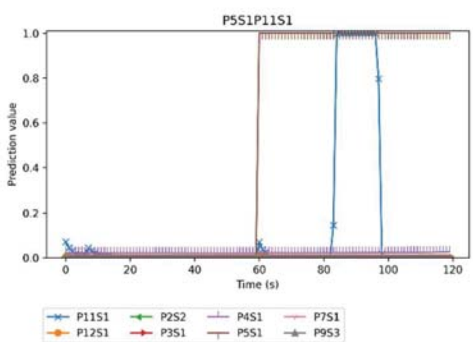

Fig. 7. Diagnosis Result of POSRV-PZR Event Scenario.

The Scenarios of Fig. 6 and Fig. 7 are when abnormal events related to each other or when abnormal events occur in the same system. Even for these cases, the proposed approach performed an accurate diagnosis. Because the cases of POSRV and PZR events occurred in proximity, the existing model showed a diagnostic accuracy of only 0.73 % in the case of simultaneous occurrence. However, the proposed approach detects both events well, despite the complex scenario in which the PZR event occurs 30 seconds later than the POSRV event.

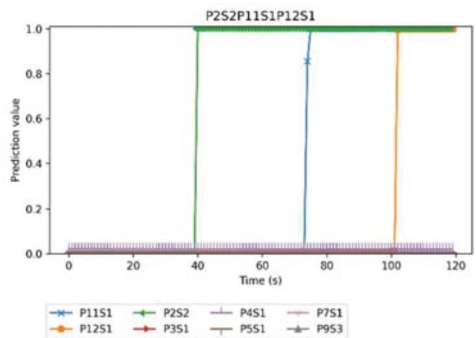

Fig. 8. Diagnosis Result of CHRG-PZR-CCW Scenario.

Above Fig. 8 is the diagnosis result of the proposed model for the case where the CHRG event, PZR event, and CCW event occur at 30-second intervals. Except for the above scenario, the other three composite abnormal event scenarios are not covered in this study. However, since the proposed approach detects each abnormal event individually, we can assume that it will have high diagnostic performance for scenarios in which two or more abnormal events, even though occurring in different times.

5. Conclusions

To apply a diagnostic model to an NPP, it should be able to show high diagnostic ability even for complex abnormal events. In this study, embedded feature selection was performed using an extra-trees classifier, and binary classification models were trained to detect each abnormal event. The model was trained with only individual abnormal event dataset and tested against complex abnormal event datasets. The proposed approach showed higher performance for scenarios in which two abnormal events occurred simultaneously compared to the general model. In addition, it detected each abnormal event even when two related abnormal events occurred at different time points. However, in the case of CHRG-POSRV, it showed lower accuracy than the general model. It can be assumed that it is because only 10 parameters were selected for detecting the POSRV event and these provided insufficient information for model diagnosis. In further study, it is necessary to perform sensitivity study on feature selection to supplement such as this point.

Acknowledgement

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (RS-2022-00144042).

References

[1] Geurts, Pierre, Damien Ernst, and Louis Wehenkel. "Extremely randomized trees." Machine learning 63 (2006): 3-42.

[2] Pereira, Rafael B., Alexandre Plastino, Bianca Zadrozny, and Luiz HC Merschmann. "Categorizing feature selection methods for multi-label classification." Artificial Intelligence Review 49 (2018): 57-78.

[3] 3KEYMASTER Simulator 2013, Western Service Corporation, Frederick, MD, USA.

[4] Nair, Vinod, and Geoffrey E. Hinton. "Rectified linear units improve restricted boltzmann machines." In Proceedings of the 27th international conference on machine learning (ICML-10), pp. 807-814. 2010.

[5] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).

[6] Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel et al. "Scikit-learn: Machine learning in Python." the Journal of machine Learning research 12 (2011): 2825-2830.