Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 9-10, 2024

Preliminary Modeling and Applicability Evaluation for Condition Diagnosis and Failure Detection in Reactor Protection System

Ho Jun Lee, Sang Hyun Lee, Man Gyun Na*
Department of Nuclear Engr., Chosun Univ., 10, Chosundae 1-gil, Dong-gu, Gwangju, 61452
**Corresponding author: magyna@chosun.ac.kr*

Keywords: LSTM, LSTM-Autoencoder, RPS, Condition diagnosis, Failure detection

1. Introduction

There are various systems in Nuclear Power Plants (NPPs) for stable operation. Among these, the role of the Instrumentation and Control (I&C) system is crucial in NPPs, where safety is a top priority. The I&C system automatically provides protection and appropriate control to manage risks that may arise in both normal and abnormal conditions. Additionally, it is responsible for initiating signals for resolution when problems occur.

Since the 2000s, these I&C systems have transitioned from analog to digital devices, driven by the obsolescence of analog systems and the benefits of easier maintenance, among other advantages. However, unlike analog devices, digital devices can breakdown instantaneously. In particular, the Reactor Protection System (RPS) among digital devices is a crucial system for safety. Such a breakdown in the RPS can lead to NPPs shutdowns and the possibility of severe accidents. According to the Operational Performance Information System for NPPs (OPIS) operated by the Korea Institute of Nuclear Safety, incidents caused by I&C defects accounted for 20.5% of all NPPs incidents from 2011 to 2023 [1]. Various efforts are being made to prevent such defects, including self-diagnosis functions, operational periodic tests, and scheduled maintenance. However, proving the reliability of self-diagnosis functions is challenging. Additionally, periodic tests occur only at scheduled times, so assessing their condition during normal operation is difficult. Furthermore, the inability to predict failures necessitates regularly replacing all components. Breakdown maintenance and preventive maintenance performed at operating NPPs are unconditional maintenance that does not consider the actual level of defects or the possibility of failure, resulting in unnecessary costs.

To address these issues, a research is underway to use Artificial Intelligence (AI) for prompt condition diagnosis and early failure prediction of RPS components, enabling proactive measures. In this paper, we present a framework for the prompt condition diagnosis and failure detection of RPS components. Additionally, by utilizing accelerated aging data from Insulated Gate Bipolar Transistor (IGBT), we conduct preliminary modeling for the condition diagnosis and failure detection of RPS components, deriving suitable AI methods.

2. Model

In this study, four AI models were utilized. Deep Neural Network (DNN) and Long Short-Term Memory (LSTM) were used in supervised learning, while Autoencoder (AE) and LSTM-AE were used for unsupervised learning.

2.1 Deep Neural Network

DNN is a type of artificial neural network featuring multiple layers. Generally, a DNN includes an input layer, several hidden layers, and an output layer. Each layer contains numerous neurons that extract features from input data and forward them to the subsequent layer. DNN excels in automatically learning features from data, making them widely used in various fields such as image recognition, natural language processing, and speech recognition.

2.2 Long Short-Term Memory

LSTM is a type of Recurrent Neural Network (RNN) designed for sequence data. It was developed to address the long-term dependency problem, a drawback of existing RNN [2]. Fig. 1 shows the structure of LSTM.

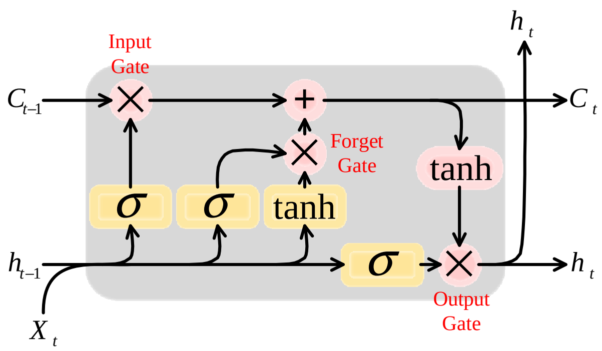

Fig. 1. LSTM structure

LSTM consists of an input gate, forget gate, and output gate. The input gate decides how much information to be added to the cell state. The forget gate determines which information to be discarded from the cell state. The output gate decides which part of the cell state to pass on as the output. This process allows LSTM to store, delete, and output the necessary information at each time step, maintaining information over long periods. Table I shows the variables of the LSTM model and their description.

Table I: LSTM Variables

Variables	Description
C_{t-1}	Cell state from the previous time step
C_t	Cell state at the current time step
X_t	Input at the current time step
h_{t-1}	Output from the previous time step
h_t	Output at the current time step

2.3 Autoencoder

AE is an artificial neural network used for efficient representation learning of data. It is primarily utilized for data dimensionality reduction, feature learning, anomaly detection, and generative models [3]. Fig. 2 shows the structure of AE.

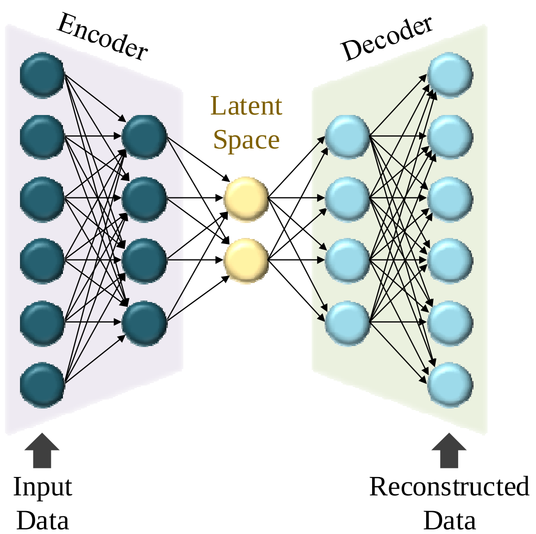

Fig. 2. AE structure

The AE is composed of two main components: the encoder and the decoder. The encoder processes input data to extract essential features, transforming it into an internal representation. The decoder takes this internal representation generated by the encoder and reconstructs data similar to the original input. The model uses a loss function to produce outputs that resemble the original data, minimizing the difference between the input and the reconstructed output. This allows the AE to capture and highlight the most significant features of the data.

2.4 Long Short-Term Memory Autoencoder

The LSTM-AE is a variation of the AE based on LSTM networks. This model aims to learn representations of sequence data efficiently. Fig. 3 shows the structure of LSTM-AE.

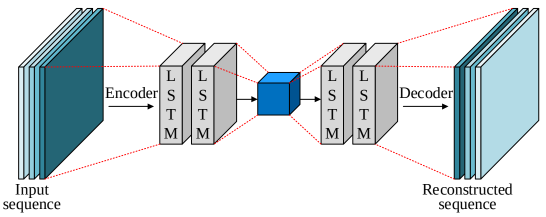

Fig. 3. LSTM-AE structure

The LSTM-AE features an architecture where both the encoder and decoder components are designed as LSTM structures. This configuration allows the model to harness the strengths of LSTM in handling sequential data. The training process of the LSTM-AE aligns with the conventional AE approach, focusing on learning efficient data representations through reconstruction. Due to the inherent properties of LSTM which excels in capturing temporal dependencies within data, the LSTM-AE is particularly well-suited for anomaly detection in time-series datasets, offering robust and effective identification of irregular patterns over time.

3. Data Preprocessing

In this study, preliminary modeling was performed using the accelerated aging data of IGBT provided by NASA's Open Data Portal [4] for condition diagnosis and failure detection. These data are from subjecting IGBT to thermal overstress aging with square signal gate voltage bias. The IGBT failure point is an instant when the latch-up phenomenon occurred. Latch-up is a phenomenon that excessive current flows due to the formation of unnecessary current paths due to strong voltage. In this study, we utilized data from four IGBT devices (Device 2, Device 3, Device 4, and Device 5). Among them, data from Device 2, Device 3, and Device 4 were used as train data for model training, and data from Device 5 was used as test data to evaluate the performance of the model. In addition, a part of the training data was selected to be used as validation data. Data preprocessing was performed to utilize this data to train an AI model. Data preprocessing involved variable selection and data standardization. Additionally, data preprocessing for the LSTM-based model was also performed using the sliding window technique. The data were divided into two sets based on the number of variables. One dataset includes three variables, while the other contains seven variables. Table II shows the input variables for two data sets. Additionally, data for each time steps of 2, 5, 10, and 15 were used using the sliding window technique.

Table II: Input variables for two data sets

Data set	Input variables
I	Operating time
	Temperature
	Voltage
II	Operating time
	Temperature
	Voltage
	Average temperature
	Average Voltage
	Weighted Average temperature
	Weighted Average Voltage

Finally, the data was divided into specific sections based on the Remaining Useful Life (RUL) to train a condition diagnosis model. Using data categorized by RUL intervals, the condition of the IGBT is diagnosed to determine which interval it falls into. Table III shows the criteria by which specific data sections were divided.

Table III: Data by specific segments based on RUL

No.	RUL
State I	RUL > 1800
State II	1300 < RUL ≤ 1800
State III	800 < RUL ≤ 1300
State IV	RUL < 800

4. Framework

Fig. 4. shows an overview of the framework for the process of condition diagnosis and failure detection.

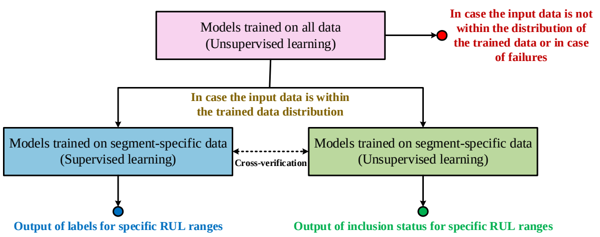

Fig. 4. Framework overview

This framework aims to construct and utilize three main models for precise diagnosis of the component conditions. The first model is an unsupervised learning model based on the entire dataset. Focused on failure detection, it identifies normal and abnormal patterns in the data, enabling early detection of unusual behaviors or potential failure states.

The second model, a supervised learning model, is based on data segmented into specific intervals. It learns from data classified according to the RUL of components and predicts condition labels for each RUL range. This allows for a more detailed understanding of component conditions.

The third model is also based on data from specific intervals but is trained using an unsupervised learning approach. Unlike the supervised model, it learns the structure and patterns of data without explicit labels to determine whether data belongs within a specific RUL interval, thereby discerning if the component condition belongs to the given RUL range.

Using these three models, we can diagnose component conditions more precisely and provide more accurate information for failure detection. Particularly, by performing dual verification of data from specific intervals through supervised and unsupervised learning models, we can enhance the predictive reliability of the models and develop a tool to support decision-making in actual operational environments.

5. Result

This study has developed failure detection and condition diagnosis models based on the designed framework. The failure detection models, AE and LSTM-AE were each constructed using the entire dataset. The LSTM-AE model with 15 time steps was designed. The output of the failure detection model is the reconstruction error value, which measures the difference between the learned data and the newly input data. By analyzing the differences in data through this reconstruction error, the model performs failure detection. The performance evaluation of the models was conducted using the Area Under the Curve (AUC) value, which represents the area under the Receiver Operating Characteristic (ROC) curve. The ROC curve is a graph that illustrates the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold levels of the model. An AUC value close to 1 indicates the superior failure detection capability of the model, while a value near 0.5 suggests that the performance is akin to random guessing. This metric serves as an important measure of how accurately the model identifies actual failure conditions.

Table IV shows the performance comparison results and shows the performance of AE and LSTM-AE according to the units and variables of the model. The LSTM-AE model with 15 time steps, 128 units and 7 variables exhibits the highest performance with an AUC value of 0.9989.

The condition diagnosis models were designed to perform cross verification using both unsupervised and supervised learning models. The condition diagnosis models, AE and LSTM-AE, were each constructed and trained using unsupervised learning. The unsupervised learning condition diagnosis model constructs a model for each specific data segment to inform which interval the current state falls into. Specific-segmented data were used as the training data. For performance evaluation, the AUC value was used as the evaluation metric.

Table V shows the results of the performance comparison, illustrating the performance of AE and LSTM-AE in specific segments based on variables. The LSTM-AE model achieved the best performance, with an AUC value of 1 across all stages, regardless of the number of variables.

The condition diagnosis models, DNN and LSTM were each constructed and trained using supervised learning. The supervised learning-based condition diagnosis model uses specifically segmented data, labeled for each state, to determine which interval the current condition falls into. The LSTM model was constructed separately for different time steps to compare performance. The metric used for performance evaluation was accuracy, which is defined as the ratio of cases correctly predicted by the model. Accuracy can be represented as Eq. (1).

$$Accuracy = \frac{\text{True Positives + True Negatives}}{\text{Total number of samples}} \quad (1)$$

Table VI shows the result of the performance comparison. The LSTM model with 15 time steps had the best performance, reaching an accuracy value of 1 across all inputs, regardless of the number of variables.

Table IV: Performance evaluation of failure detection model using AUC

Model	Units	AUC	
		3 Inputs	7 Inputs
AE	16	0.9791	0.9295
	32	0.9723	0.9404
	64	0.9564	0.8105
	128	0.913	0.8724
	256	0.9826	0.7956
	512	0.9675	0.8557
LSTM-AE (15 time steps)	32	0.9973	0.9962
	64	0.9761	0.9971
	128	**0.9981**	**0.9989**
	256	0.9757	0.9844

Table V: Performance evaluation of unsupervised learning condition diagnosis model using AUC

Model	State	AUC	
		3 Inputs	7 Inputs
AE	State I	0.9817	0.9484
	State II	0.9642	0.9579
	State III	0.9503	0.9857
	State IV	0.9707	0.9981
LSTM-AE (15 time steps)	**State I**	**1.0**	**1.0**
	State II	**1.0**	**1.0**
	State III	**1.0**	**1.0**
	State IV	**1.0**	**1.0**

Table VI: Performance evaluation of supervised learning condition diagnosis model using accuracy

Model	Accuracy	
	3 Inputs	7 Inputs
DNN	0.8312	0.8961
LSTM (2 time steps)	0.8356	0.8496
LSTM (5 time steps)	0.8361	0.8525
LSTM (10 time steps)	0.9268	0.9756
LSTM (15 time steps)	**1.0**	**1.0**

6. Conclusions

In this study, an AI-based framework is proposed for condition diagnosis and failure prediction of the RPS components, using the accelerated aging data of IGBT provided by NASA for initial modeling. This research was conducted using four AI models: DNN, LSTM, AE, and LSTM-AE. This study aims to accurately diagnose the condition of RPS components and predict failures early, thereby improving the safety of NPPs.

Experimental results showed that the LSTM-AE model with 128 units, 7 variables, and 15 time steps exhibited the best performance among all the fault detection models built using the entire dataset, with an AUC value of 0.9989. This indicates that the LSTM-AE model is capable of accurately capturing the temporal continuity of data, and based on this, it can precisely detect failure.

Furthermore, in the case of unsupervised learning models for condition diagnosis, the LSTM-AE model showed the best performance with an AUC of 1 when trained on segment specific data. In the case of supervised learning models, the LSTM model using 15 time steps showed the best performance with an accuracy of 1. This suggests that LSTM can effectively handle long-term dependencies in data and accurately judge the condition of RPS components.

Through this study, it is expected that the AI-based diagnostic and prediction models to be developed in the future will accurately assess the condition of RPS components in NPPs and detect potential failures early, contributing to the safe operation of NPPs. These models are also expected to provide advantages such as improving the efficiency of regular maintenance work and reducing unnecessary replacement costs.

Acknowledgment

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government (MOTIE) (20224B10100120, Development of commercialization technology for failure diagnosis of reactor control and digital I&C systems) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2022-00144239).

REFERENCES

[1] Operational Performance Information System for Nuclear Power Plant, "Nuclear Accident and Failure Status", 2023 last modified Dec 13, https://opis.kins.re.kr/accident/CauseAccidentFailPresent.do

[2] S. Hochreiter and J. Schmidhuber, Long Short-Term Memory, in Neural Computation, Vol.9, pp.1735-1780, 1997.

[3] M. A. Kramer, Nonlinear principal component analysis using autoassociative neural networks, ALCHE J. 37 (2) (1991) 233-243.

[4] J. Celaya, P. Wyscocki, and K. Goebel, IGBT Accelerated Aging Data Set, NASA Prognostics Data Repository, NASA Ames Research Center, Moffett Field, CA, 2009.