Pertanika J. Sci. & Technol. 34 (4): 2053 - 2074 (2026)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Data-Driven Fault Detection and Diagnosis of a Nuclear Reactor Cooling System Using Graph Neural Networks

Dhines Shree Kanthan[1], Nur Aina Mohammad Aziz[1], Azura Che Soh[1]*, Ribhan Zafira Abdul Rahman[1], Mohd Khair Hassan[1], Mohd Sabri Minhat[2], and Julia Abdul Karim[2]

[1]*Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*

[2]*Reactor Instrumentation and Control Engineering, Reactor Technology Centre, Malaysian Nuclear Agency (Nuklear Malaysia), 43000 Kajang, Selangor, Malaysia*

ABSTRACT

The cooling system of a nuclear reactor is essential for ensuring the safety and efficient operation of the TRIGA PUSPATI Reactor (RTP). However, conventional fault detection methods often struggle with the nonlinearities and complexity of the system. This study proposes a data-driven Fault Detection and Diagnosis (FDD) framework for the RTP cooling system using a Graph Neural Network (GNN). In this approach, the system is modelled as a graph, where sensors are represented as nodes and their interactions as edges, enabling the GNN to capture component relationships and improve anomaly detection accuracy. A scaled-down control rig was developed to replicate the RTP cooling system for experimental data collection and validation. The model was trained and evaluated using experimental data, and its performance was assessed using evaluation metrics including accuracy, precision, recall, and F1 score. For comparison, an Artificial Neural Network (ANN) model was also implemented. The results show that the GNN outperforms the ANN in fault detection performance. Both models were evaluated under offline conditions, while real-time validation was conducted for the GNN. In the online implementation, a Wi Fi module was used to transmit live sensor data to the cloud for continuous monitoring. The control rig operates independently in a laboratory environment and is not connected to any operational nuclear reactor system, with cloud communication used strictly for experimental monitoring without control actions. The results demonstrate that the proposed GNN achieves accurate and reliable

ARTICLE INFO

Article history:
Received: 15 April 2026
Accepted: 13 July 2026
Published: 13 August 2026

DOI: https://doi.org/10.47836/pjst.34.4.05

E-mail addresses:
gs75025@student.upm.edu.my (Dhines Shree Kanthan)
gs69737@student.upm.edu.my (Nur Aina Mohammad Abdul Aziz)
azuracs@upm.edu.my (Azura Che Soh)
ribhan@upm.edu.my (Ribhan Zafira Abdul Rahman)
khair@upm.edu.my (Mohd Khair Hassan)
sabri@nm.gov.my (Mohd Sabri Minhat)
julia@nm.gov.my (Julia Abdul Karim)
* Corresponding author

e-ISSN: 2231-8526

fault detection in both offline and real-time applications, highlighting its potential to enhance reactor safety and operational reliability.

Keywords: Artificial neural network, data-driven, fault detection and diagnosis, graph neural network, reactor cooling system

INTRODUCTION

State-of-the-art fault detection and diagnosis (FDD) methods can generally be divided into two main groups: traditional or physics-based techniques and Artificial Intelligence (AI)-based approaches. Statistical Process Control (SPC), model-based and rule-based systems are typical approaches (Bersimis et al., 2007; Martinez-Viol et al., 2020; Qian et al., 2024; Simiyon et al., 2024). Granted, SPC applies statistical techniques such as control charts to control process variables; however, it is based on clearly defined and simple processes (Bersimis et al., 2007). The model-based approach uses mathematical models to compare model circumstances or prediction anomalies between expectation and actual behaviour, but it encounters difficulties when applied to nonlinear, uncertain and complex systems (Simiyon et al., 2024). Rule-based approaches rely on expert-design threshold and are inflexible to accommodate more sophisticated systems if the task is getting harder (Qian et al., 2024).

In recent years, advancements in machine learning and data analytics, driven by the availability of big data, have brought significant attention towards data-driven techniques (Bon & Van Dai, 2022). Among these, Artificial Neural Networks (ANNs) are widely adopted for their capability to capture complex patterns for classification and anomaly detection tasks, but it requires large datasets (de Albuquerque Filho et al., 2022; Pang et al., 2021). Alternatively, Support Vector Machines (SVMs) can achieve strong classification performance, but they require careful hyperparameter tuning and may become computationally expensive while handling large datasets (Bharathwaj et al., 2025; de Souza et al., 2014; Liu et al., 2021).

Furthermore, deep learning techniques such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) can select hierarchical features from raw time-series data, eliminating the need for manual feature engineering. Despite this advantage, their practical implementation is often constrained by the requirement for large datasets of labelled data and more powerful computational resources due to their complex architectures and numerous parameters (Mienye, 2024; Kong et al., 2023).

In contrast, Graph Neural Networks (GNNs) represent a more recent and expressive paradigm for systems that consist of interacting elements. Due to processing graph-structure data with nodes representing components and edges indicating relationships, GNN can capture dependencies not captured by traditional neural networks (Hoang et al., 2023; Khemani et al., 2024; Theiler & Fink, 2024). This capability makes them particularly well-suited for the demanding and complex conditions encountered in industrial systems.

This research develops a GNN-based FDD system for complex industrial cooling systems. A scaled-down process control rig replicates the RTP cooling system to safely generate operational data under both normal and fault conditions, which cannot be performed directly on the reactor. The system is modelled to capture its dynamics and interactions, and a GNN-based FDD algorithm is developed and validated on the control rig platform, efficiently learning complex interdependencies and providing adaptive, data-driven monitoring. This approach offers a safe, scalable and intelligent solution for monitoring complex interdependent systems and advancing fault detection research in industrial settings.

Process Control Rig for RTP Cooling System

The RTP cooling is based on the components which are designed to remove the heat from the reactor in a structured manner. The cooling system passes off the heat from the reactor to its tower, which emits it into the atmosphere. Critical sensors such as temperature, pressure and flow measure system status always. The first pump passes the coolant from the reactor to a cooling tower, while the second is responsible for circulating water in a heat exchanger. In each of the loops, flow and pressure are controlled by valves to keep the process at safe operating conditions. Scaled-down process control rig and its P&ID diagram are provided in Figures 1 and 2, respectively.

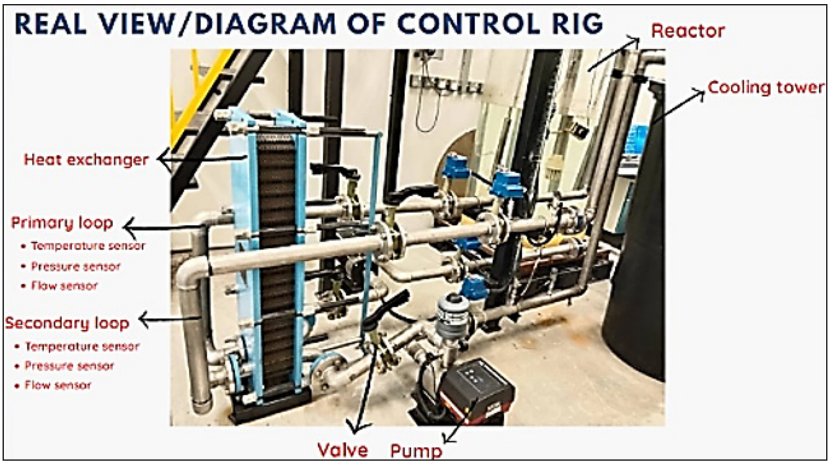

Figure 1. Real view of the process control rig for the RTP cooling system

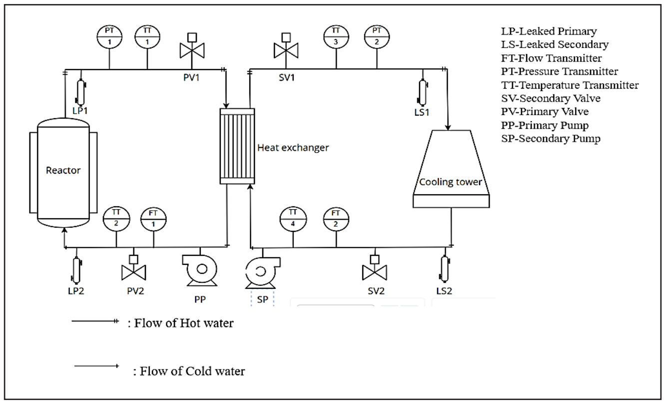

Figure 2. P&ID diagram of process control rig for RTP cooling system

The P&ID shows the plumbing of sensors, and flow paths: double line is for hot water pipelines and Single lines are for cold water pipes. It also displays relationships among the reactor, heat exchanger, pumps, valves and sensors, as well as with a cooling tower to explain how heat travels and is manipulated throughout components in the system.

The pilot plant control rig was developed based on the actual RTP cooling system to support experimental data acquisition under normal and fault conditions for FDD development. The rig is designed as a scaled-down functional validation platform with an approximate 1:20 overall water volume ratio and is not intended to reproduce exact reactor thermal hydraulic behaviour, but rather to emulate operational logic, control system response, and degradation behaviour of structures, systems, and components (SSCs). Although the system architecture of the RTP cooling system is preserved through a simplified configuration of primary and secondary loops, including tanks, pumps, a heat exchanger, and an electric heater assembly, local fluid dynamic phenomena cannot be fully replicated due to scaling effects.

Thermal-hydraulic distortions are therefore unavoidable, particularly in natural circulation effects, heat transfer coefficients, and flow distribution characteristics. The electric heater (1.5 kW total, consisting of three 500 W elements) provides a controlled heat source to represent reactor core loading in a qualitative manner for control system evaluation rather than nuclear heat reproduction. The piping system is simplified to 2-inch diameter lines compared to the actual 4-inch (primary) and 6-inch (secondary) RTP system, and demineralised water is not used, introducing minor deviations in thermophysical properties. Instrumentation includes RTD PT100 sensors, Type K thermocouples, flow meters, and pressure sensors for system-level monitoring and control validation. Despite these limitations, the test rig provides a safe and effective platform for control development, educational training, and AI-based fault diagnosis studies, enabling macro-level analysis of system behaviour and fault scenarios without risk to actual reactor operation.

METHODOLOGY

A project of project for this study starts with gathering data from the system. The FDD system model is established, and the GNN model for FDD is proposed online and offline tested accordingly. If sufficient performance of the model is achieved, it is implemented to perform analysis on the results. A GUI is then generated to provide user access to the system. If the model does not satisfy the requirement, it will continue to optimise its parameters and validate again.

Experiment and Data Collection

This section describes the data collection procedure for the RTP cooling system process control rig. The system configuration and sensor setup are summarised in Table 1, while eight fault scenarios representing the degradation of critical system components are presented in Table 2. Before fault injection, the cooling system was operated under normal conditions to establish a stable baseline, and normal operating data were recorded. For each experimental run, the system was allowed to reach steady-state operation before data acquisition to ensure that the recorded measurements accurately represented system behaviour under both normal and fault conditions. A sampling duration of 20 minutes was adopted for each operating condition.

The dataset was collected using six sensors installed on the RTP cooling system control rig, comprising two pressure transmitters (PT1 and PT2), three temperature transmitters (TT1, TT2, and TT3), and one flow transmitter (FT1). Sensor measurements were acquired at a sampling rate of 1 Hz, corresponding to one sample per second throughout each experiment. The complete dataset contained 28,644 samples covering both normal operation and the eight fault scenarios.

In this study, a GNN model is proposed for modelling the RTP cooling system, while a conventional ANN model is used as the benchmarking method for performance comparison. The dataset was randomly divided into 70% training, 15% validation, and 15% testing subsets. The training set was used to optimise the model parameters, while the validation set monitored the generalisation performance and reduced the risk of overfitting during training. The independent testing set was reserved exclusively for the final performance evaluation. Cross-validation was not employed because a fixed training, validation, and testing partition was adopted to ensure a consistent and fair comparison between the proposed GNN model and the benchmark ANN model. In addition, weight decay regularisation and validation monitoring were incorporated to further enhance the generalisation capability of both models.

Table 1
Types of components with their functions and fault scenarios

Components	Functions	Fault Scenarios
Primary Valve 1 (PV1)	Regulates coolant flow within the primary loop	Failure of PV1
Primary Valve 2 (PV2)	Controls coolant flow in the primary cooling system	Failure of PV2
Leaked Primary Pump 1 (LP1)	Ensures circulation of coolant in the reactor's primary loop; any leaks can result in reduced flow	Failure of LP1
Leaked Primary Pump 2 (LP2)	Ensures coolant circulation in the primary loop to absorb core heat; leaks in PV2 can reduce flow.	Failure of LP2
Leaked Secondary Pump 1 (LS1)	The secondary loop transfers heat to the cooling tower; leaks reduce coolant flow	Failure of LS1
Leaked Secondary Pump 2 (LS2)	Ensures coolant circulation in the secondary loop; leaks in S2 may decrease the flow rate.	Failure of LS2

Table 2
Types of faults and their injection method

Fault	Fault Type	Fault Injection
1	Restricted water flow rate in the Hot-Heat Exchanger & Reactor	Incremental closure of PV1: 30%, 60%, and 90%
2	Restricted water flow rate in the Hot-Heat Exchanger & Cooling Tower	Incremental closure of SV1: 30%, 60%, and 90%
3	Restricted water flow rate in the Hot-Heat Exchanger & Cooling Tower	Increase the opening of LS1 from 50% to 100%
4	Restricted water flow rate in the Cold-Heat Exchanger & Reactor	Incremental closure of PV2: 30%, 60%, and 90%
5	Restricted water flow rate in the Cold-Heat Exchanger & Reactor	The PP was turned OFF
6	Restricted water flow rate in the Cold-Heat Exchanger & Cooling tower	Increase the opening of LS2 from 50% to 100%
7	Restricted water flow rate in the Cold-Heat Exchanger & Cooling tower	Incremental closure of SV2: 30%, 60%, and 90%
8	Restricted water flow rate in the Cold-Heat Exchanger & Cooling tower	The SP was turned OFF

Modelling of RTP Cooling System

Figure 3 presents the bock diagram of RTP cooling system, including the relationship between inputs and outputs, such as the main transmitter that indicates how the different system components correlate to achieve proper thermal regulation and operational stability.

The cooling system depends on Temperature Transmitters (TT), Pressure Transmitters (PT) and Flow Transmitters (FT) to control temperature, flow and pressure inside the reactor loops. The incoming coolant flow and temperature are measured by FT2 and TT4 in the cold heat exchanger, while PT2 and TT3 measure the outgoing pressure and temperature to ascertain a good heat transfer. These outputs, along with reactor PT1 and TT1, become inputs for the hot heat exchanger, whose outputs FT1 and TT2 are fed back to the reactor, providing a control means relating to cooling ability and safety. This feedback mechanism permits the system to respond to changes and achieve a desired level of thermal and pressure conditions in the reactor core.

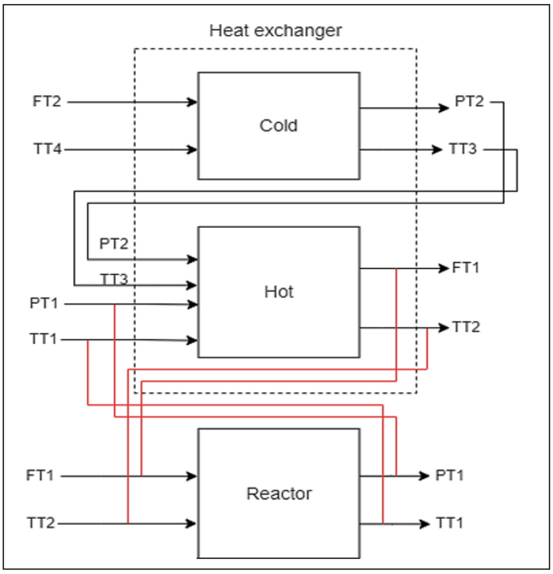

Figure. 3 Block diagram of RTP cooling system

GNN-Based RTP Cooling System Model

The model architecture of the GNN including three layers (input, subsystem, and output), is depicted in Figure 4. The input layer is used to take the temperature (TT), flow (FT), and pressure (PT) sensor readings from various locations in the cooling circuit. For example, FT2 and TT4 are for the Cold Heat Exchanger. While PT2, TT3, PT1 and TT1 are for the Hot Heat Exchanger and FT1 and TT2 are for the Reactor. In contrast, the output layer is used as both input for the next subsystem and as the feedback mechanism. A node is represented by each subsystem while connecting its interactions.

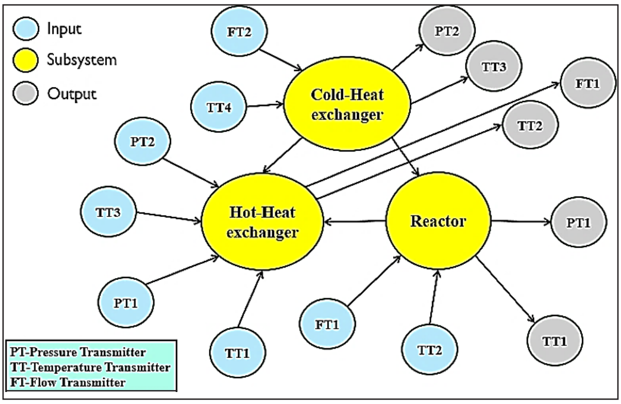

Figure 4. Structure of the GNN model for RTP cooling system control rig

In GNNs, several hyperparameters may impact the model's accuracy, including performance, convergence rate, and generalisation capability. Inaccurate configuration can cause overfitting, underfitting, slow convergence, or high computational cost. Therefore, these parameters should be properly calibrated to ensure stable and reliable model performance.

The number of hidden layers (N_H) may influence the depth of the GNN and affect its capability to capture the nonlinearities in the system. While deeper architectures can improve the model's behaviour, an excessive number of hidden layers may introduce training instability and performance reduction. Prior studies indicate that GNNs generally achieve optimal performance with two to four hidden layers, whereas performance tends to decline when the number exceeds seven layers (Kipf & Welling, 2017; Xu et al., 2019). Nevertheless, when labelled data or node features are limited, deeper architectures may be beneficial to support semi-supervised learning.

The hidden layer dimension (N_{HD}) determines the size of the node embeddings and showcases the model's ability to represent complex features. For relatively simple tasks such as node classification, embedding dimensions of 16, 32, or 64 are typically sufficient (Liu et al., 2021). Increasing the embedding dimension can improve the model's representation capability but also increase computational cost, whereas smaller dimensions may decrease effective feature extraction.

In comparison to conventional neural networks, GNNs consider information from neighbouring nodes using a graph structure. Thus, the number of neighbouring nodes (N_{NN}) is a critical parameter, as it represents how much surrounding information is related during message transfer. Larger embedding dimensions often require a broader neighbourhood to capture meaningful relationships between nodes. Previous research indicates that an optimal number of neighbouring nodes generally lies between 5 and 50, as values outside this range can decline the model performance (Chen et al., 2023).

The number of samples being processed in each training cycle is decided based on the batch size (B_S), It significantly affects both convergence behaviour and memory usage. Smaller batch sizes allow faster parameter updates and are suitable for simpler models, one that is used in RTP cooling systems, where values between 1 and 32 are often effective (Hu et al., 2023). On the other hand, larger batch sizes can improve training stability but require more computational resources.

The number of graph convolution layers (N_{CL}) is another key architectural parameter, particularly in convolution-based GNNs. Multiple factors are considered, including the number of hidden layers, embedding dimensions, neighbouring nodes, and batch size. Typically, the optimal number of layers ranges from one to four, as deeper networks are prone to over-smoothing, where node features become overly similar (Zhang et al., 2023).

The number of training iterations (N_T) directly impacts training efficiency. Excessive iterations can increase computational cost and lead to overfitting, while too few may result in underfitting. Studies suggested that optimal iteration counts typically range between 50 and 1000, depending on dataset size and model complexity.

Finally, the learning rate also plays a crucial role in controlling convergence speed and stability. Commonly adopted values range from 0.01 to 1, particularly when adaptive or decaying learning strategies are employed (Yang et al., 2023). An excessively high learning rate may cause divergence, whereas an overly small value can slow convergence and increase the risk of getting stuck in local minima.

The forward propagation process in a GNN updates node features through successive graph convolution layers according to the formulation introduced in the Graph Convolutional Network framework, as expressed in Equation 1 (Kipf & Welling, 2017):

$$H^{l+1} = \sigma\left(W^{(l)} H^{(l)} A + B^{(l)}\right) \qquad [1]$$

where $H^{(l)}$ is the feature matrix at layer l, $W^{(l)}$ and $B^{(l)}$ are trainable weights and biases, respectively, A is the adjacency matrix representing node connections, and σ is a non-linear activation function such as ReLU.

Model parameters are optimised using gradient descent, which updates the parameters iteratively according to the standard optimisation rule widely used in neural networks, as expressed in Equation 2 (Rumelhart et al., 1986):

$$\theta^{(t+1)} = \theta^{(t)} - l_r \cdot \nabla_\theta J(\theta) \qquad [2]$$

where θ represents the model parameters, l_r is the learning rate, and $J(\theta)$ is the loss function.

To improve generalisation and mitigate overfitting, weight decay is incorporated through L2 regularisation, as commonly adopted in machine learning optimisation, as expressed in Equation 3 (Goodfellow et al., 2016):

$$J(\theta) = L(\theta) + W_d \sum_i \theta_i^2 \qquad [3]$$

where $J(\theta)$ is the original loss function and W_d is the weight decay coefficient.

The sigmoid parameter (S_p) controls the nonlinearity of the sigmoid activation function, originally introduced in classical neural network models, as expressed in Equation 4 (Rumelhart et al., 1986):

$$(x) = \frac{1}{1 + e^{-S_p x}} \quad [4]$$

where S_p controls the steepness of the activation function and directly influences the model's sensitivity to input variations. If S_p is too large, the function saturates rapidly, leading to vanishing gradients during training. Conversely, a very small S_p may slow down the learning process and hinder convergence.

The model is trained using datasets which include temperature, pressure, and flow rate measurements collected under both normal operations and simulated fault scenarios. During training, the model aims to minimise the Mean Squared Error (MSE), which is a metrics used to validate regression learning that compute the average squared difference between actual and predicted values, as expressed in Equation 5 (Hastie et al., 2009).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad [5]$$

where y_i represents the actual value, $\hat{y}_i$ is the predicted value, and n denotes the total number of samples.

Lower MSE values, together with regression coefficients approaching one, indicate that the GNN effectively captures the hidden pattern within the RTP cooling system. This enables the model to distinguish between normal and faulty operating conditions, thereby supporting accurate fault detection and classification. Table 3 presents a summary of the final GNN architecture and selected hyperparameter settings, including the number of layers, activation functions, learning rate, batch size, and training epochs.

Table 3
GNN modelling with parameter settings

Hyperparameter	Symbol	Value range
Hidden Layers	N_h	2
Hidden Layer Dimension	N_{HD}	90
Number of Neighbor Nodes	N_{NN}	5
Batch Size	B_S	32
Graph Convolution Layers	N_{CL}	3

Table 3 (*continued*)

Hyperparameter	Symbol	Value range
Number of Training	N_t	100
Learning Rates	l_r	0.01
Weight Decay	λ	0.0001
Sigmoid Parameter	S_P	1.0

ANN Benchmark Model

Figure 5 illustrates the architecture of the Artificial Neural Network (ANN), representing the conventional machine learning (ML) model used for modelling the RTP cooling system. The ANN employs a feedforward neural network (FNN) architecture comprising an input layer, two fully connected hidden layers, and an output layer. The input layer receives sensor measurements from the RTP cooling system, including pressure transmitters (PT), temperature transmitters (TT), and flow transmitters (FT). These input features are propagated through two hidden layers, each consisting of 90 neurons with the Rectified Linear Unit (ReLU) activation function, enabling the network to learn the nonlinear relationships between the process variables and the corresponding fault classes.

The output layer employs a linear activation function to predict the target outputs associated with the monitored cooling system variables. The ANN model was trained using the Adam optimiser with a learning rate of 0.01, a batch size of 32, and 100 training epochs. Model optimisation was performed by minimising the MSE loss function through the backpropagation algorithm. These hyperparameter settings were selected to ensure a fair comparison with the proposed GNN, allowing both models to be trained under identical experimental conditions.

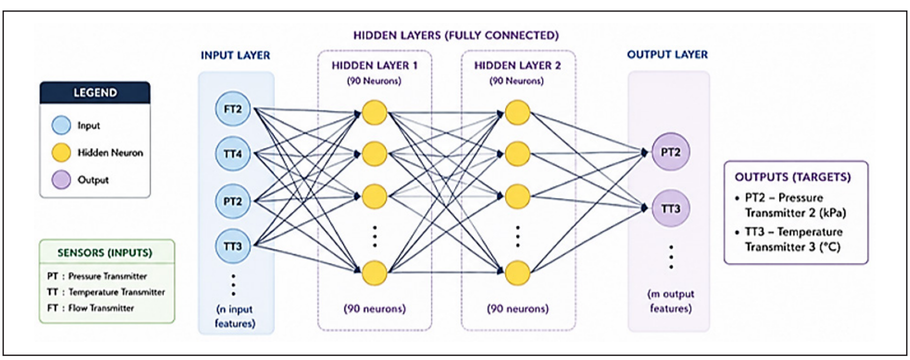

Figure 5. Structure of the ANN model for the RTP cooling system control rig

Unlike the proposed GNN, which explicitly represents the RTP cooling system as a graph of interconnected sensors, the conventional ANN processes each sensor measurement as an independent input feature through fully connected layers. Consequently, the ANN is unable to directly capture the physical relationships and fault propagation among pressure, temperature, and flow sensors. This limitation reduces its capability to model complex subsystem interactions, leading to lower fault classification performance compared with the graph-based GNN. Nevertheless, the ANN serves as a widely accepted benchmark for evaluating the effectiveness of the proposed GNN architecture in industrial FDD applications. Table 4 summarises the ANN architecture and selected hyperparameter settings.

Table 4
ANN modelling with parameter settings

Hyperparameter	Symbol	Value range
Hidden Layers	N_h	2
Hidden Layer Dimension	N_{HD}	90
Number of Neighbor Nodes	N_{NN}	-
Batch Size	B_S	32
Graph Convolution Layers	N_{CL}	-
Number of Training	N_t	100
Learning Rates	l_r	0.01
Weight Decay	λ	0.0001
Sigmoid Parameter	S_P	1.0

Verification of Offline and Online GNN Approaches for FDD System

In this work, only the proposed GNN model is evaluated under both offline and online testing scenarios. The GNN (Offline) solution is trained using historical data, including temperature, pressure, and flow rate measurements under normal reactor operation conditions. When the GNN (Offline) model is trained, its predictions are considered as a baseline for comparison against the performance of GNN (Online).

Apart from the Offline GNN models, the Online GNN operates in real-time using live data via Wi-Fi adapter from the cloud and logging sensor readings are uploaded automatically to Google Drive from the cooling system's control rig directly, as described in Figure 6. This enables the system to operate continuously and respond to any fault conditions within the reactor cooling circuit. After the real-time data is processed by the GNN (Online) model, it is stored in a server and made available for system analysis and review.

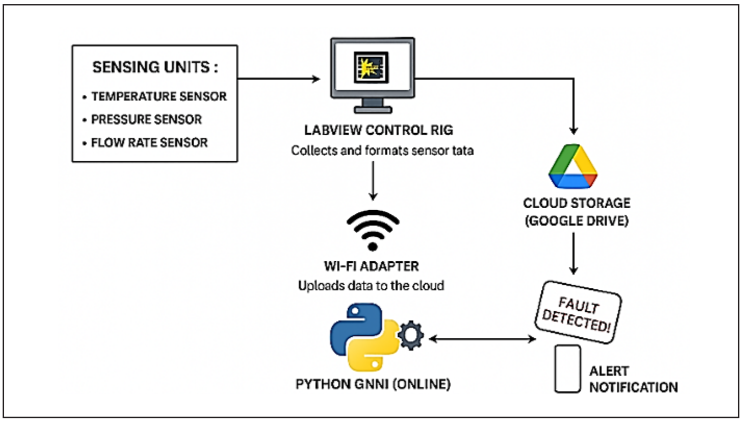

Figure 6. Online FDD implementation using GNN on the RTP cooling system control rig

To verify the validity and information reliability of our GNN (Online) method, we compare online real-time prediction results with offline GNN-based models. Such a comparative method can also verify the performance of the GNN (Online) model and make sure it makes predictions likely to the history. This validation ensures the correctness of the real-time model and thus increases fault detection reliability.

It is important to note that the pilot plant control rig operates independently within a laboratory research environment and is not connected to any operational nuclear reactor system. The wireless communication and cloud storage are used strictly for experimental monitoring and research purposes, and the system does not perform any control actions. For practical deployment in nuclear facilities, the architecture would require adaptation to comply with cybersecurity requirements, such as air-gapped or segmented networks, secure plant-level communication infrastructure, data diode-based unidirectional transmission, and edge computing deployment for local processing without reliance on external cloud services. These approaches ensure that the conceptual monitoring framework demonstrated in the test rig can be translated into a secure, plant-compliant architecture without violating nuclear cybersecurity regulations.

RESULTS AND DISCUSSIONS

This section presents the modelling results of the RTP cooling system using the proposed GNN model, with the ANN serving as the benchmark model. Both models are evaluated under normal operating conditions and faulty scenarios in the offline analysis. For the online analysis, only the proposed GNN model is evaluated. The FDD results are then presented, followed by an overall evaluation of the proposed system's performance.

RTP Cooling System Modelling and Analysis

The proposed GNN provides an accurate representation of the RTP cooling system under normal operating conditions. The model was trained using experimentally observed data collected from the RTP cooling system and successfully captured the dynamic behaviour of the three major subsystems, namely the cold-heat exchanger, hot-heat exchanger, and reactor. To further evaluate the effectiveness of the proposed approach, its performance was compared with a conventional ANN trained using the same dataset and identical training conditions.

Model performance was evaluated using the mean residual value. A residual below 10% was classified as Rank A, indicating excellent agreement between the predicted and measured system responses. Residuals greater than 10% but below 30% were classified as Rank B, residuals exceeding 50% were classified as Rank D, and all remaining residuals were classified as Rank C (Atkinson et al., 2017). As presented in Table 5, the proposed GNN achieved Rank A for all subsystem outputs, with mean residual values ranging from 0.05% to 1.75%, demonstrating excellent prediction accuracy across all three subsystems. In contrast, although the conventional ANN could model the overall system behaviour, it produced consistently higher residual values. The ANN achieved Rank A for PT2, FT1, and PT1, but only Rank B for TT3, TT2, and TT1, indicating lower prediction accuracy for the temperature-related outputs.

Table 5
Comparative performance evaluation of the RTP cooling system subsystem model using the proposed GNN and the conventional ANN

Subsystem	Output	Mean Residual (%) GNN	Ranking (GNN)	Mean Residual (%) ANN	Ranking (ANN)
Cold-Heat Exchanger	PT2	1.75	A	7.84	A
	TT3	0.06	A	12.65	B
Hot-Heat Exchanger	FT1	0.16	A	8.21	A
	TT2	0.06	A	11.43	B
Reactor	PT1	1.75	A	9.38	A
	TT1	0.05	A	10.72	B

Note. A = good; B = acceptable; C = marginal; D = poor

Table 6 presents the evaluation metrics comparison between the proposed GNN and the conventional ANN for FDD of the RTP cooling system. Both models were trained and evaluated using the same experimentally collected dataset and identical training conditions to ensure a fair comparison. The proposed GNN achieved 98.12% Accuracy, 99.15% Precision, 99.26% Recall, and 98.20% F1-score, whereas the conventional ANN achieved 96.68% Accuracy, 96.72% Precision, 96.63% Recall, and 96.67% F1-score.

These results indicate that the proposed GNN successfully classified all normal and fault operating conditions without misclassification, while the ANN exhibited a small number of incorrect classifications.

In addition to the classification metrics, the proposed GNN demonstrated superior regression performance by achieving an MSE of 0.0002, compared with 0.0184 for the conventional ANN. Similarly, the GNN produced a lower average residual value of 1.70%, whereas the ANN recorded an average residual of 3.18%. Furthermore, the proposed GNN achieved a classification error of 0.01%, while the ANN produced a classification error of 3.32%. These findings demonstrate that the GNN provides more accurate predictions of the pressure, temperature, and flow sensor responses under both normal and faulty operating conditions.

The superior performance of the proposed GNN is attributed to its graph-based learning architecture, which explicitly models the physical relationships among interconnected pressure, temperature, and flow sensors within the RTP cooling system. By learning both the sensor features and their structural dependencies, the GNN effectively captures complex system dynamics and fault propagation, resulting in lower prediction errors and higher modelling accuracy than the conventional ANN. In contrast, the ANN treats each sensor measurement as an independent feature and cannot explicitly exploit the interdependencies among sensors, limiting its ability to represent the underlying system behaviour. Consequently, the proposed GNN provides a more robust and reliable framework for fault detection and diagnosis, demonstrating that graph-based learning is better suited for modelling complex industrial cooling systems than conventional ANN approaches.

Table 6
Evaluation metrics comparison between the proposed GNN and the conventional ANN

Performance Metric	Proposed GNN	Conventional ANN
Accuracy (%)	98.12	96.68
Precision (%)	99.15	96.72
Recall (%)	99.26	96.63
F1-score (%)	98.20	96.67
MSE	0.0002	0.0184
Average Residual (%)	1.70	3.18
Classification Error (%)	0.01	3.32

Performance Evaluation of the GNN-Based FDD System in Offline Operation

An FDD algorithm based on GNN has been proposed for the system in normal conditions and under some faulty conditions. The offline and online validation of the algorithm in both cases was investigated experimentally. Below shows the corresponding values to indicate its capability of distinguishing normal from faulty states.

Under offline mode, the processed data are initially uploaded to the GNN-based FDD system to be analysed, as shown in Figure 7. The uploaded data is pre-processed and compared with the baseline normal operating conditions that are learned by the GNN model. The operational condition of the reactor control rig is determined in accordance with this calculation.

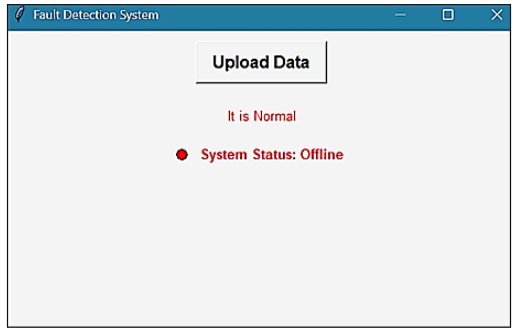

Figure 7. Offline GNN-based FDD under normal conditions

The final system status and fault location are visualised through the GNN representation depicted in Figure 8. In this diagram, the colour of the node denotes the state of the system; for instance, green nodes represent normal operation. Such a graphical representation has been considered to offer an intuitive and efficient method of monitoring system operation and detecting possible fault locations in offline modes as well.

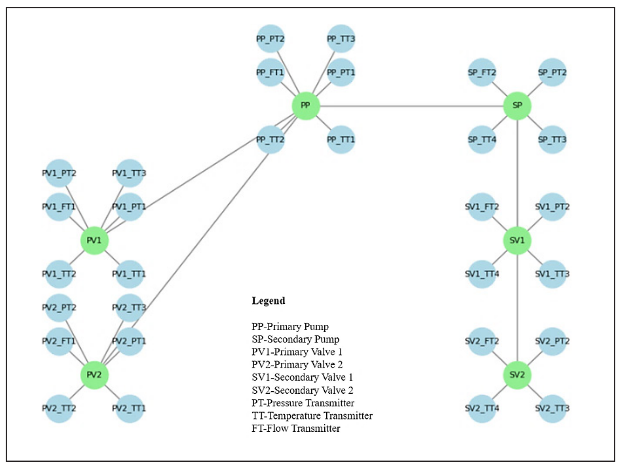

Figure 8. Offline GNN-based FDD: location mapping under normal conditions

Several fault scenarios were tested offline with actual fault data from Primary Pump (PP) and Secondary Valve 1 (SV1). When the measured data is uploaded to the system, the GNN mode l with detect and diagnose the faults by comparison of the measurements with normal operating conditions. The tendencies of the condition in PP and SV1 are shown in Figures 9 and 10, respectively.

The identified faults are represented by red-coloured nodes in the visualisations of Figures 11 and 12 to illustrate that the system can provide evidence for faults occurring in PP and SV1. These results demonstrate that the system is capable of accurate fault detection and able to provide direct indication of both the type and location of faults in the RTP cooling system rig.

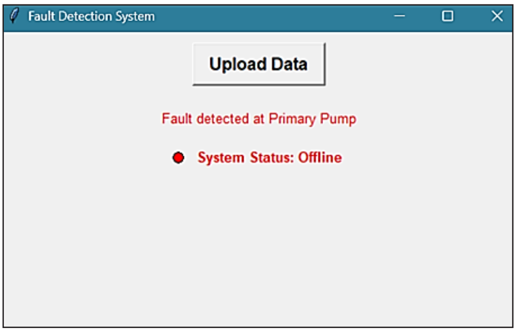

Figure 9. An offline GNN-based FDD system identifies a fault in the primary pump

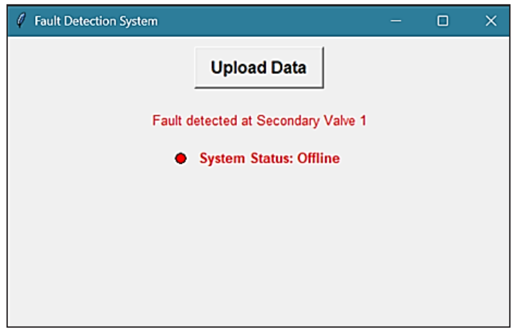

Figure 10. An offline GNN-based FDD system identifies a fault in the secondary valve 1

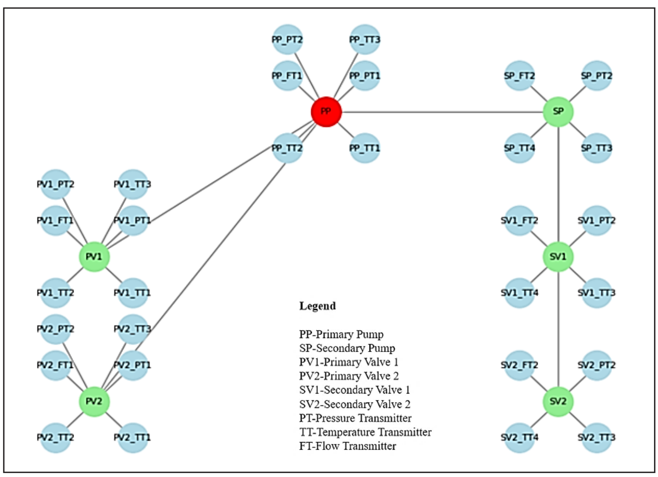

Figure 11. Offline GNN-based FDD system for fault localisation at the primary pump

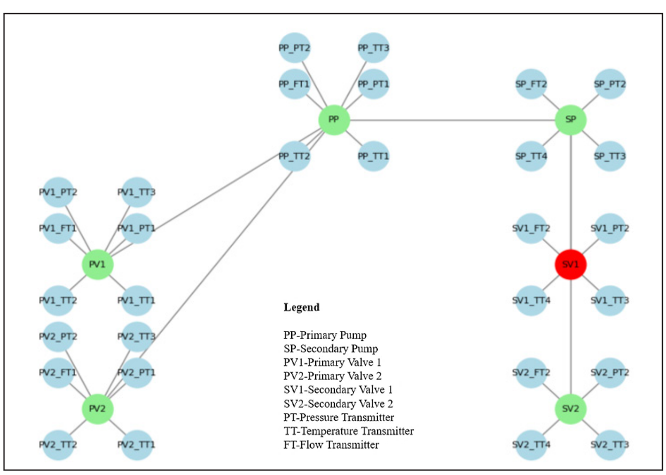

Figure 12. Offline GNN-based FDD system for fault localisation at the secondary valve 1

Performance Evaluation of the GNN-Based FDD System in Online Operation

The GNN-based FDD system was additionally validated via online analysis by interfacing it with a real-time implementation. This arrangement ensured that the system could also actively detect and locate faults as they arose. During this process, the system was operating in a normal manner at first. For the online validation of the fault scenarios tested, faults were injected one by one while running and re-running the above scenarios with a focus on PP, and SV1 was done as only single faults were tested.

The online validation agreed well with the offline analysis, indicating that faults have been consistently and accurately detected with both methods. These verifications of faults for the PP and SV1 are depicted in Figures 13 and 14. Furthermore, the types and location of faults are exactly those which can be seen in Figures 15 and 16 for the presented visualisations in Figures 11 and 12. These validation results demonstrate the robustness and reliability of the proposed GNN-based FDD system, highlighting its strong potential for real-time implementation in the actual RTP cooling system, as presented in the confusion matrix shown in Figure 17.

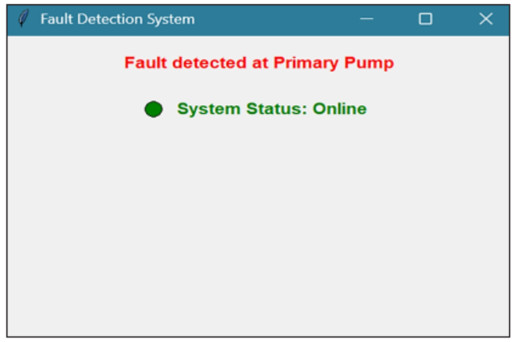

Figure 13. An online GNN-based FDD system identifies a fault in the primary pump

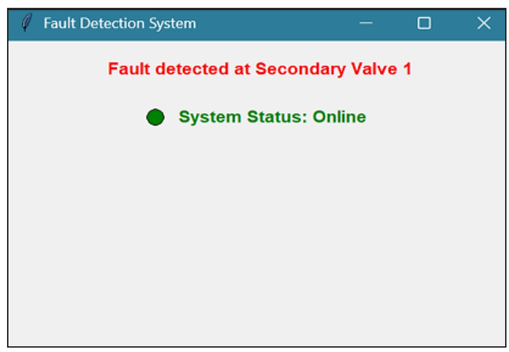

Figure 14. An online GNN-based FDD system identifies a fault in the secondary valve 1

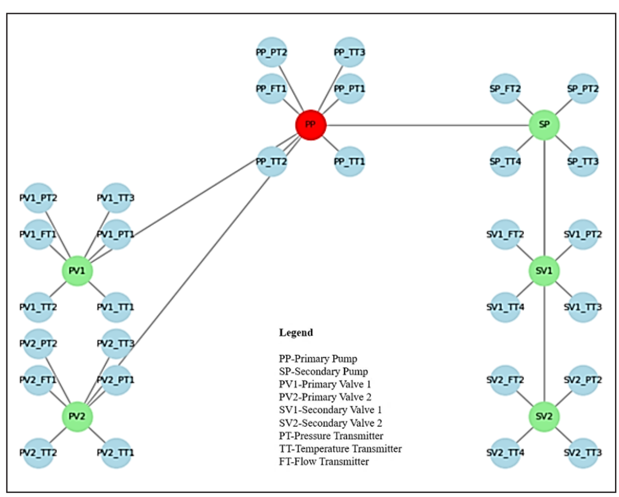

Figure 15. Online GNN-based FDD system for fault localisation at the primary pump

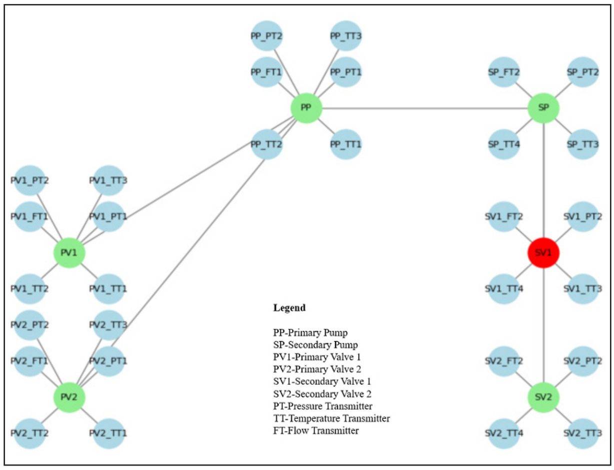

Figure 16. Online GNN-based FDD system for fault localisation at the secondary valve 1

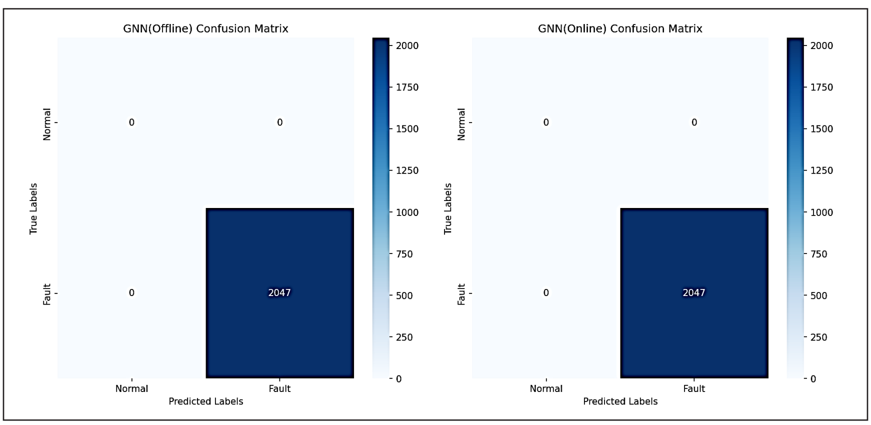

Figure 17. Confusion matrix of GNN (offline) and GNN (online)

CONCLUSIONS

GNN-based FDD algorithms were developed and validated offline for normal operation and chosen fault cases with Primary Pump as well as Secondary Valve 1. Through comparison of incoming data with a predetermined baseline, the system was able to effectively discriminate normal and faulty conditions and automatically determine the fault mode. Such an automatic procedure allows a faster identification and localisation of failures and can contribute to the system's improved reliability.

Online real-time tests were then conducted on a test rig applying the developed GNN-based FDD system, connected to components of the cooling system downstream of the pipeline. Online results were consistent with the offline evaluation, showing that models based on traditional baselines could well generalise to real operational scenarios.

These results also prove the effectiveness, precision and reliability of the GNN method as a fault detection and predictive maintenance solution for the RTP cooling system.

In summary, the proposed study has successfully designed an intelligent FDD system for a reactor cooling system of nuclear power plants with GNN. The key goals of the research were the construction of a GNN fault detection model, the creation of an intuitive graphical user interface, and proof-of-concept validation using data generated from the cooling system process control rig. All objectives were achieved successfully. Therefore, we can achieve accurate FDD for invasive faults in offline and online scenarios and significantly improve the maintenance efficiency and operational safety directly.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the Ministry of Science, Technology and Innovation (MOSTI) for providing funding to support the development of the RTP cooling system rig platform for this research. The authors also express their sincere thanks to the Malaysian Nuclear Agency for their guidance during the experimental work, assistance with data collection, and technical advice in testing the algorithms, which were instrumental to the successful completion of this research.

Data and Code Availability Statement

The data and code used in this study are not publicly available due to confidentiality restrictions.

REFERENCES

Atkinson, A. D., Hill, R. R., Pignatiello, J. J., Vining, G. G., White, E. D., & Chicken, E. (2017). Wavelet ANOVA approach to model validation. *Simulation Modelling Practice and Theory, 78*, 18-27. https://doi.org/10.1016/j.simpat.2017.08.004

Bersimis, S., Psarakis, S., & Panaretos, J. (2007). Multivariate statistical process control charts: An overview. *Quality and Reliability Engineering International, 23*(5), 517-543. https://doi.org/10.1002/qre.829

Bharathwaj, G., Khurshid, A., Venkatavijayan, S., & Pani, A. K. (2025). Industrial process monitoring using support vector data description: A systematic review and application for fault detection in multiphase flow systems. *Flow Measurement and Instrumentation, 106*, Article 103031. https://doi.org/10.1016/j.flowmeasinst.2025.103031

Bon, N. N., & Van Dai, L. (2022). Fault identification, classification, and location on transmission lines using combined machine learning methods. *International Journal of Engineering and Technology Innovation, 12*(2), 91-109. https://doi.org/10.46604/ijeti.2022.7571

Chen, S., Wulamu, A., Zou, Q., Zheng, H., Wen, L., Guo, X., Chen, H., Zhang, T., & Zhang, Y. (2023). MD-GNN: A mechanism for data-driven graph neural networks for molecular properties prediction and new material discovery. *Journal of Molecular Graphics and Modelling, 123*, Article 108506. https://doi.org/10.1016/j.jmgm.2023.108506 de Albuquerque Filho, J. E., Brandão, L. P., Fernandes, B. J. T., & Maciel, A. M. A. (2022). A review of neural networks for anomaly detection. *IEEE Access, 10*, 112342-112367. https://doi.org/10.1109/ACCESS.2022.3216007

de Souza, D. L., Granzotto, M. H., de Almeida, G. M., & Oliveira Lopes, L. C. (2014). Fault detection and diagnosis using support vector machines: A SVC and SVR comparison. *Journal of Safety Engineering, 3*(1), 18-29. https://doi.org/10.5923/j.safety.20140301.03

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. https://doi.org/10.7551/mitpress/10243.001.0001

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Hoang, V. T., Jeon, H.-J., You, E.-S., Yoon, Y., & Jung, S. (2023). Graph representation learning and its applications: A survey. *Sensors, 23*(8), Article 4168. https://doi.org/10.3390/s23084168

Hu, B., Zhang, Y., Liu, Z., & Wang, J. (2023). Adaptive batch size selection for efficient deep neural network training. *IEEE Transactions on Neural Networks and Learning Systems, 34*(11), 5481-5493. https://doi.org/10.1109/TNNLS.2023.3272941

Khemani, B., Patil, S., Kotecha, K., & Tanwar, S. (2024). A review of graph neural networks: Concepts, architectures, techniques, challenges, datasets, applications, and future directions. *Journal of Big Data, 11*(1), Article 18. https://doi.org/10.1186/s40537-023-00876-4

Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *Proceedings of the 5th International Conference on Learning Representations (ICLR 2017)*. https://doi.org/10.48550/arXiv.1609.02907

Kong, X., Chen, Z., Liu, W., Ning, K., Zhang, L., Marier, S. M., Liu, Y., Chen, Y., & Xia, F. (2025). Deep learning for time series forecasting: A survey. *International Journal of Machine Learning and Cybernetics, 16*, 5079–5112. https://doi.org/10.1007/s13042-025-02560-w

Liu, C., Cichoń, A., Królczyk, G., Li, Z., & Zhu, X. (2022). Technology development and commercial applications of industrial fault diagnosis systems: A review. *The International Journal of Advanced Manufacturing Technology, 118*(11-12), 3497-3529. https://doi.org/10.1007/s00170-021-08047-6

Martinez-Viol, V., Urbano, E. M., Prieto, M. D., & Romeral, L. (2020). HVAC early fault detection using a fuzzy logic-based approach. *Renewable Energy and Power Quality Journal, 18*(2), 196-201. https://doi.org/10.24084/repqj18.270

Mienye, I. D. (2024). A comprehensive review of deep learning: Architectures, algorithms and limitations. *Information, 15*(12), Article 755. https://doi.org/10.3390/info15120755

Pang, G., Shen, C., Cao, L., & van den Hengel, A. (2021). Deep learning for anomaly detection: A review. *ACM Computing Surveys, 54*(2), Article 38. https://doi.org/10.1145/3439950

Qian, H., Pan, Y., Wang, X., & Li, Z. (2024). Research on the optimisation of belief rule bases using the Naive Bayes theory. *Frontiers in Energy Research, 12*, Article 1396841. https://doi.org/10.3389/fenrg.2024.1396841

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*(6088), 533-536. https://doi.org/10.1038/323533a0

Simiyon, A., Sachidanand, C., Halmakki Krishnamurthy, M., Bhatt, A. V., & Indiran, T. (2024). Machine learning-based fault classification in a pilot plant batch reactor using a support vector machine. *ACS Omega, 9*(26), 29041-29052. https://doi.org/10.1021/acsomega.4c04421

Theiler, R., & Fink, O. (2024). Graph neural networks for electric and hydraulic data fusion to enhance short-term forecasting of pumped storage hydroelectricity. In *Proceedings of the European Conference of the PHM Society, 8*(1), 1-11. https://doi.org/10.36001/phme.2024.v8i1.4129

Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2019). How powerful are graph neural networks? In *Proceedings of the 7th International Conference on Learning Representations (ICLR 2019)*. https://doi.org/10.48550/arXiv.1810.00826

Yang, Z., Liu, Z., Zhou, J., Song, C., Xiang, Q., He, Q., Hu, J., Faber, M. H., Zio, E., Li, Z., Su, H., & Zhan, J. (2023). A graph neural network (GNN) method for assigning gas calorific values to natural gas pipeline networks. *Energy, 278*, Article 127875. https://doi.org/10.1016/j.energy.2023.127875

Zhang, H., Nguyen, D. H., & Tsuda, K. (2023). Differentiable optimisation layers enhance GNN-based mitosis detection. *Scientific Reports, 13*, Article 4306. https://doi.org/10.1038/s41598-023-41562-y