The Scientific Temper (2023) Vol. 14 (1): 206-210
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2023.14.1.25
https://scientifictemper.com/

RESEARCH ARTICLE

LSTM based data driven fault detection and isolation in small modular reactors

Swetha Rajkumar, Jayaprasanth Devakumar

Abstract

Nuclear power stations revealed their value in the power sector by supplying reliable, emission-free power for many years. The highest standards of safety must be attained since a nuclear power station is a nonlinear, intricate, time-varying system that has the probability of leaking radiations. Pr edominantly, it is challenging for operators to quickly and precisely extract critical data about the real plant variables as a result of the vast monitoring data obtained in modern NPPs. However, current developments in machine learning techniques have made it conceivable for operators to interpret these vast amounts of data and take appropriate action. Thermal hydraulic analysis using the RELAP5 algorithm was done on the IP-200 NPP. A long short-term memory architecture was trained to categorize six different simulated IP-200 circumstances. The outcomes improved the accuracy and dependability of nuclear power plant fault monitoring systems.

Keywords: Deep learning, Fault detection and isolation, Long short-term memory, Pressurized water reactor, Recurrent neural network, Small modular reactor.

Introduction

Since they are safety-critical systems, early fault identification is compulsory for pressurized water reactors. This can be done by using analytical redundancy components (Kumar *et al.*, 2022) or process data analytics (Kumar and Devi (2021). The nuclear industry is continually working to construct safer and more effective reactors for its upcoming generation of plants. To help operators make the right findings in the case of an anomaly or problem, thereby raising the level of safety for these reactors, is one of the essential conditions for attaining that goal. In this regard, fault monitoring techniques may significantly contribute to raising these plants' safety standards. The inherent shortcomings of conventional methods have made intelligent fault diagnosis procedures a lively area of research (Ma and Jiang, 2009). Based on the recent trend, machine learning is currently the most required option for future fault diagnosis technology (Patan, 2008)

Reported recent studies in automated fault diagnosis include: comparing anomaly diagnosis ability between a Radial Basis Architecture and an Elman neural architecture (Kumar, 2021). Ayodeji, Liu and Xia (2018) proposed a fault diagnosis scheme based on Support Vector Machines (SVM) that is skilled of component-level tasks. In paper by Peng *et al.* (2018), the Sequential Probability Ratio Assessment was used to spot faults in PWR by means of distributed hierarchy.

This article proposes a data-based anomaly diagnosis procedure built on the long short-term memory (LSTM) architecture. Conventional neural algorithms are susceptible to overfitting and frequently peak in local minima. Furthermore, deciding on the architecture structure, with the number of hidden neurons, is tough. These create practical issues when attempting to implement the conventional vanilla neural architecture as a resolution. recurrent neural architectures (RNN) was developed for sequence problems to address this issue. This method outperforms conventional ANN and is better suited for structuring time series data (Zhang, Wang and Liu, 2008; Zhang *et al.*, 2010; Zhang *et al.*, 2013; Kumar and Devakumar, 2022). It's likewise remained used to aid in anomaly diagnosis.

Conventional RNN generates nonlinear models for detecting actuator faults (Talebi, Khorasani and Tafazoli, 2009). Furthermore, the Elman architecture's convergence speed and generalization ability have been improved.

Department of Instrumentation and Control Systems Engineering, PSG College of Technology, Coimbatore, Tamil Nadu, India

***Corresponding Author:** Swetha Rajkumar, Department of Instrumentation and Control Systems Engineering, PSG College of Technology, Coimbatore, Tamil Nadu, India, E-Mail: srk.ice@psgtech.ac.in

How to cite this article: Kumar, S.R., Jayaprasanth, D. (2023). LSTM based data driven fault detection and isolation in small modular reactors. The Scientific Temper, **14**(1):206-210

Doi: 10.58414/SCIENTIFICTEMPER.2023.14.1.25

Source of support: Nil

Conflict of interest: None.

Received: 11/01/2023 **Accepted:** 12/02/2023 **Published** : 25/03/2023

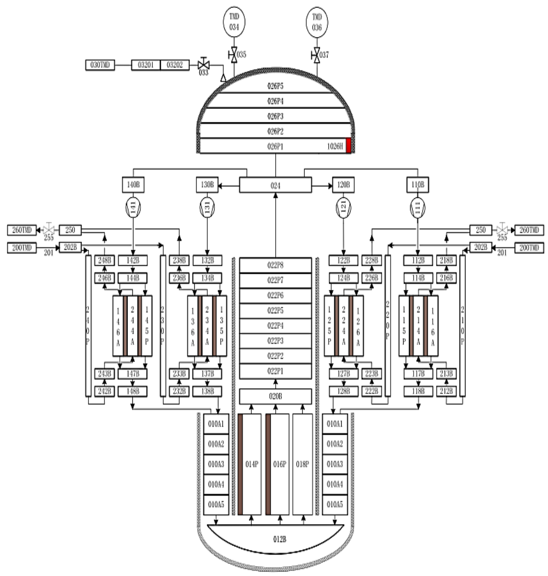

Figure 1: Nodalization of IP-200

Though, if RNN permits for information storing, architecture gradients incline to fade over time. LSTM architecture employs RNNs. It is particularly good at speech recognition, text classification, and other tasks. In the field of nonlinear systems, anomaly prediction is regarded as the primary subject (Wu *et al*., 2018). Until now, slight research on using LSTM in anomaly diagnosis are carried out. Depending on its outcome in time series data, the investigation goals to implement LSTM for anomaly diagnosis in a small modular reactor.

The content discussed is as follows. Section 2 briefs on the normalization and parameters of small modular reactors, fault scenarios, and LSTM. Section 3 explains on the results obtained. Section 4 provides the conclusion and future scope.

Materials and Methods

IP-200 - Modular Reactor

The IP-200, a cutting-edge small modular reactor (SMR) depending upon Integral PWR modeled at Harbin Engineering University, is under consideration. Medium and small-sized reactors (SMR) have grown in acceptance during recent years owing to their small size, the comfort of assembly (at the plant), and approachability in distant locations (Locatelli *et al*., 2014). This type of reactor is perfect for powering distant and seaside areas. The reactor's main parameters are revealed in Table 1. RELAP5 software was used to model the thermal hydraulics of the IP-200 reactor (Saeed *et al*., 2020; Wang *et al*., 2021; Kumar & Devakumar, 2022). This is a two-fluid scheme code for analysis. Figure 1 depicts the system's normalization.

Due to its integral design, the pressure container has the majority of the primary system, with the core, pressurizer,

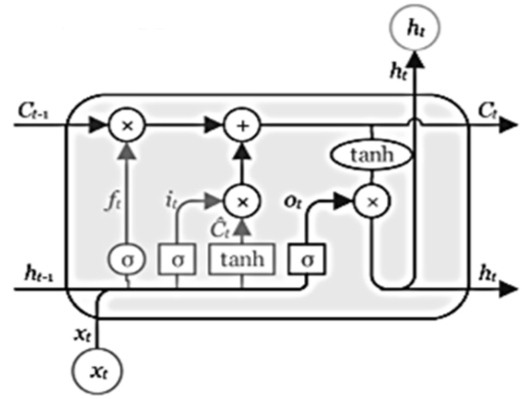

Figure 2: LSTM schematic layer

OTSG (Once Through Boilers/Steam Generators), and pumps. (Xia *et al*., 2016; Sun *et al*., 2017; Jiang *et al*., 2018) Labels its functioning. The reactor core contains plate category fuel elements, which is situated at the lowest of the reactor container. It involves three flow conduits: a hot (014P) channel, a bypass (018P) channel, and an average (016P) channel. Though the upper plenary functions as a pressurizer with safety features built in. Four clusters of OSTGs are utilized for heat transmission to the secondary unit, each one linked to a pump located in the superior side of the reactor vessel. Shared headers serve both the feed water (202B) inlet and the superheated steam (250) outlet.

Fault Scenarios

The IP-200 RELAP-5 simulated dataset can be found at (Saeed *et al*., 2020; Wang *et al*., 2021; Kumar & Devakumar, 2022). To estimate the efficacy of the proposed anomaly diagnosis structure, a learning was carried out using the thermal hydraulic guesstimate code in RELAP5 for simulation of the IP-200 NPP. The various abnormal and normal simulated circumstances considered in this learning are enumerated in Table 2.

A Long Short-term Memory (LSTM) architecture is utilized in this study to classify the above-simulated conditions.

Table 1: Key Parameters of IP-200

Process Parameters	*Rated Values*
Power at full core	220 MW
Pressurizer Pressure	15.5 MPa
Inlet Temperature of Reactor Core	562.15 K
Outlet Temperature of Reactor Core	594.15 K
Mass flow rate of Feed water	81.5 kg/s
Temperature of Feed water	373.15 K
Mass flow rate of Primary Coolant	1200 kg/s
Superheat of steam	40.0 K
Pressure of steam	3.0 MPa
Number of Once through Steam Generators	12
Number of main pumps	4

Table 2: Simulated conditions with values

Classification Number	*Simulated Condition*
0	Steady state at rated Power
1	Steam generator tube rupture
2	Reactor Coolant Pump failure
3	Pressurizer PORV struck open 100%
4	Power transient 60 to 80%
5	Feedwater line break 50%

Long Short-Term Memory

Recurrent neural architecture (RNN) is an upgraded type of traditional neural architecture. The temporal statistics is the input information, with contacts amongst units forming a directed cycle in that layer. In distinction, a traditional neural architecture has connections in between the layers only. Inside a single layer, there is no linking between the units. Thus, this architecture does not diffuse time series numbers, its effectiveness for time series numbers will be subpar. But, if RNNs can stock data over time, the gradients in the architecture vanishes. If it happens, the learning capability of the RNN reduces significantly.

The LSTM neural architecture is categorized as deep-net, which is intended to acquire long-term dependencies. It can preserve data for lengthy epochs of time by utilizing the introduced gates. The forget gate is cast off to ignore repeated data, the input gate chooses critical data to be saved in the internal state, and the output gate predicts output data.

Hochreiter and Schmidhuber devised LSTM network (Hochreiter and Schmidhuber, 1997), which Gers refined and advanced (Gers, Schmidhuber and Cummins, 2000). While the horizontal line in the normal RNN neural architecture layer passes over the uppermost of the diagram, the LSTM contains three gates to defend and govern the cell state. The first stage of the LSTM is to determine which data from the cell state will be ignored. The choice was made using a method known as forgetting the door. ht-1 and xt is read from the gate for each quantity in the cell state and outputs a value between 0 and 1. Ct-1. 1 denotes "fully reserved," while 0 denotes "totally discarded."

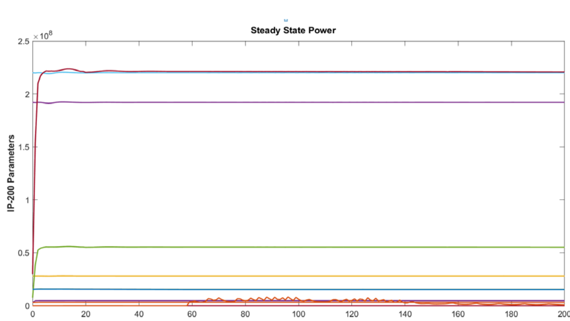

Figure 3: Steady state at rated Power

Table 3: LSTM Training parameters

Layers	LSTM LSTM LSTM Fully Connected SoftMax Classification
Number of inputs	43 sensor signals from IP-200
Number of classifications	6
Gradient Threshold	0.001
Maximum Epochs	150
Number of neurons	Hidden layer 1: 150 Hidden layer 2: 150 Hidden layer 3: 150

The following step determines the new data saved in the cell state. This article is alienated into two parts. The sigmoid layer, also known as the 'input layer,' resolves which value to inform first. Then, using a tanh layer, a next candidate vector is created, and Ct is supplementary to state. In the subsequent stage, these two messages will provide status information. Multiply the present state by ft and remove any unnecessary data. Then mix it with Ct. This will be the most recent candidate, and it deviations depending on how much each state is altered. Finally, decide on a output value. The output is based on the cells' present status, then it will be filtered.

The sigmoid layer is used to regulate portion of the cell's state that will be produced as the first step. The cell's state is then multiplied by the sigmoid gate output using tanh (a value between -1 and 1), with just the portion that regulates the output being produced. The equations of the LSTM are

"*Forget gate:* $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ (1)

Input gate, new cell: $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$ (2)

$c_t' = tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$ (3)

Update cell: $c_t = f_t * c_{t-1} + i_t * c_t'$ (4)

Output gate: $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ (5)

$h_t = o_t \cdot tanh(c_t)$ "(6)

Figure 2 shows a quick schematic of the LSTM layer. The projected LSTM architecture has an optimizer that functions on the SGDM. The classic Stochastic Gradient Descent (SGD) procedure changes architecture parameters (biases and weights) to minimize the objective function by taking incremental stages in the direction of the loss function's negative incline.

"$\theta_{l+1} = \theta_l - \alpha \nabla E(\theta_l)$ (7)

where l epitomizes epochs, α>0 the rate of learning, remains the parameter vector, and E() the objective function. The incline of t this objective function, ∇E(θ), is premeditated by means of the whole training data. The conventional gradient descent technique usages all data in a single pass.

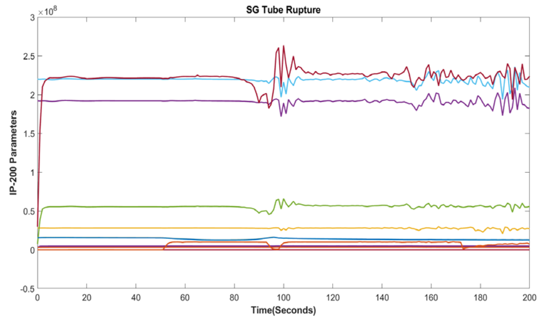

Figure 4: Steam generator tube rupture

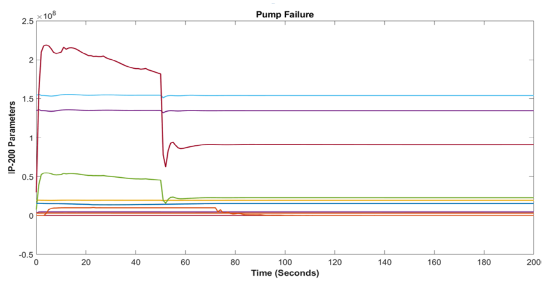

Figure 5: Reactor Coolant Pump failure

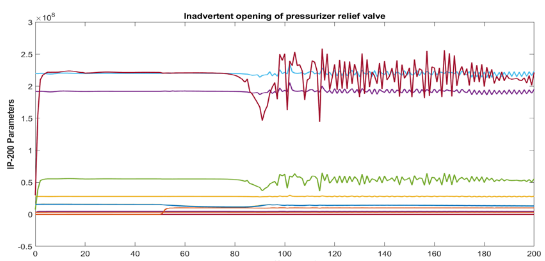

Figure 6: Pressurizer PORV struck open 100%

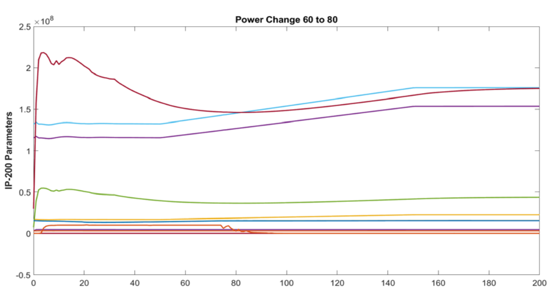

Figure 7: Power transient 60 to 80%

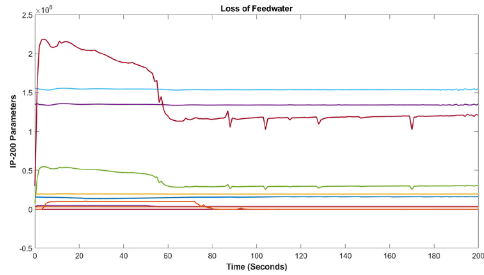

Figure 8: Feedwater line break 50%

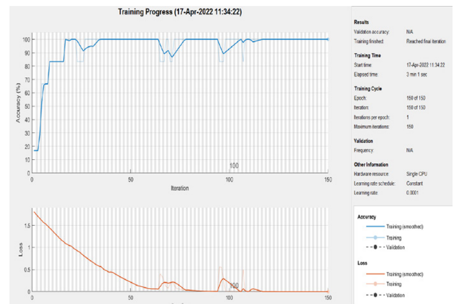

Figure 9: Architecture accuracy and loss progress

The SGD method might fluctuate through the steepest descent route as it approaches the optima. One method for reducing fluctuation is to include a momentum term in the parameter update.

The SGDM update is

$$\theta_{l+1} = \theta_l - \alpha \nabla E(\theta_l) + \gamma(\theta_l - \theta_{l-1}) \tag{8}$$

where the influence of the preceding gradient stage to the present epoch is determined one method for reducing overfitting is to include a regularisation term in the weights of the objective function, E(). The regularisation term is also known as weight decay. With the regularisation term, the loss function looks like:

$$E_R(\theta) = E(\theta) + \lambda\Omega(w) \tag{9}$$

where λ - the regularization factor, w - the weight vector, and the function of regularization   (w) is

$$\Omega(w) = 1/2w^T w'' \tag{10}$$

The LSTM neural architecture can handle classification architecture through recognized input and fault categories. This technique is praised in industries for a variability of tasks. (20)

Results and Discussion

The Relap-5 dataset was analyzed through Figures 3 to 8. Simulated sensor data from 43 process variables like flow, temperature, and pressure of several elements of IP-200 is considered under study. The plots indicate the variable changes due to the presence of fault.

A LSTM architecture was trained in MATLAB environment for classifying IP-200 simulated fault and normal conditions. The parameters used for training are as in Table 3

Six layers of neural architectures was trained for classification via 43 different process parameters. The process parameter value differs for each simulated fault/non-fault condition. Thus, the architecture was trained using six sets of 43 process parameters. The number of neurons in the hidden layers was chosen in such a way that nominal performance accuracy is achieved. The training performance of LSTM is as shown in Figure 9.

It is noted that the trained LSTM model can accurately classify the six types of simulated conditions with identical efficiency for all faults.

Conclusion

An online fault monitoring technique for a nuclear power plant is presented in this work. The LSTM architecture reliably classifies four fault conditions: Steam generator tube rupture, reactor coolant pump failure, pressurizer valve open, and feedwater line break, as well as two non-fault conditions: power transient and steady-state at rated power. The IP-200 Relap-5 dataset was used in this investigation. The trained LSTM architecture appropriately classifies the above six cases. Other fault and non-fault scenarios will be included in future studies, with the goal of using LSTM to forecast process variables and to diagnose and isolate faults. Other deep learning algorithms can be investigated and their performance compared to that of LSTM.

References

Ayodeji, A., Liu, Y. K., & Xia, H. (2018). Knowledge base operator support system for nuclear power plant fault diagnosis. Progress in nuclear energy, 105, 42-50.

Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. Neural computation, 12(10), 2451-2471.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

Jiang, N., Peng, M., & Cong, T. (2018). Simulation analysis of an open natural circulation for the passive residual heat removal in IPWR. Annals of Nuclear Energy, 117, 223-233.

Kumar, S. R., & Devakumar, J. (2022). Performance evaluation of neural network topologies for online state estimation and fault detection in pressurized water reactor. Annals of Nuclear Energy, 175, 109235.

Kumar, S. R., & Gayathri, R. G. (2021, December). Supervised machine leaning based anomaly detection and diagnosis in grid connected photovoltaic systems. In Proceedings of the First International Conference on Combinatorial and Optimization, ICCAP 2021, December 7-8 2021, Chennai, India.

Kumar, S. R., Iniyal, U. S., Harshitha, V., Abinaya, M., Janani, J., & Jayaprasanth, D. (2022, March). Anomaly Detection in Centrifugal Pumps Using Model Based Approach. In 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 427-433). IEEE.

Kumar, S.R. (2021) 'Fault Classification in Boiler Drum Using SVM and KNN Prediction Algorithms', International Journal of Mechanical Engineering, 6(3), 7.

Locatelli, G., Bingham, C., & Mancini, M. (2014). Small modular reactors: A comprehensive overview of their economics and strategic aspects. Progress in Nuclear Energy, 73, 75-85.

Ma, J., & Jiang, J. (2009). Applications of fault diagnosis in nuclear power plants: an introductory survey. IFAC Proceedings Volumes, 42(8), 1150-1161.

Patan, K., 2008. Artificial Neural Networks for the Modelling and Fault Diagnosis of Technical Processes. Springer

Peng, B. S., Xia, H., Liu, Y. K., Yang, B., Guo, D., & Zhu, S. M. (2018). Research on intelligent fault diagnosis method for nuclear power plant based on correlation analysis and deep belief network. Progress in Nuclear Energy, 108, 419-427.

Saeed, H. A., Wang, H., Peng, M., Hussain, A., & Nawaz, A. (2020). Online fault monitoring based on deep neural network & sliding window technique. Progress in Nuclear Energy, 121, 103236.

Sun, L., Peng, M., Xia, G., Lv, X., & Li, R. (2017). Numerical study on coolant flow distribution at the core inlet for an integral pressurized water reactor. Nuclear Engineering and Technology, 49(1), 71-81.

Talebi, H. A., Khorasani, K., & Tafazoli, S. (2008). A recurrent neural-network-based sensor and actuator fault detection and isolation for nonlinear systems with application to the satellite's attitude control subsystem. IEEE Transactions on Neural Networks, 20(1), 45-60.

V. Peddinti, Y. Wang, D. Povey, S. Khudanpur, Low latency acoustic modeling using temporal convolution and LSTMs, IEEE Signal Process. Lett. 25 (2018) 373–377

Wang, H., Peng, M. J., Ayodeji, A., Xia, H., Wang, X. K., & Li, Z. K. (2021). Advanced fault diagnosis method for nuclear power plant based on convolutional gated recurrent network and enhanced particle swarm optimization. Annals of Nuclear Energy, 151, 107934.

Wu, Y., Yuan, M., Dong, S., Lin, L., & Liu, Y. (2018). Remaining useful life estimation of engineered systems using vanilla LSTM neural networks. Neurocomputing, 275, 167-179.

Xia, G. L., Su, G. H., & Peng, M. J. (2016). Analysis of natural circulation operational characteristics for integrated pressurized water reactor. Annals of Nuclear Energy, 92, 304-311.

Zhang, H., Liu, Z., Huang, G. B., & Wang, Z. (2009). Novel weighting-delay-based stability criteria for recurrent neural networks with time-varying delay. IEEE Transactions on Neural Networks, 21(1), 91-106.

Zhang, H., Wang, Z., & Liu, D. (2008). Global asymptotic stability of recurrent neural networks with multiple time-varying delays. IEEE Transactions on Neural Networks, 19(5), 855-873.

Zhang, H., Yang, F., Liu, X., & Zhang, Q. (2013). Stability analysis for neural networks with time-varying delay based on quadratic convex combination. IEEE transactions on neural networks and learning systems, 24(4), 513-521.