

Research Article

The Research of Fault Diagnosis of Nuclear Power Plant Based on ELM-AdaBoost.SAMME

Cheng Li ^{1,2}, Ren Yu ¹ and Tianshu Wang ¹

¹Naval University of Engineering, Wuhan 430033, China

²China Nuclear Power Operation Technology Corporation, Ltd., Wuhan 430000, China

Correspondence should be addressed to Cheng Li; 296379210@qq.com

Received 27 October 2020; Accepted 10 December 2020; Published 21 December 2020

Academic Editor: Han Zhang

Copyright © 2020 Cheng Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A fault diagnosis framework based on extreme learning machine (ELM) and AdaBoost.SAMME is proposed in a nuclear power plant (NPP) in this paper. After briefly describing the principles of ELM and AdaBoost.SAMME algorithm, the fault diagnosis framework sets ELM algorithm as the weak classifier and then integrates several weak classifiers into a strong one using the AdaBoost.SAMME algorithm. Furthermore, some experiments are put forward for the setting of two algorithms. The results of simulation experiments on the HPR1000 simulator show that the combined method has higher precision and faster speed by improving the performance of weak classifiers compared to the BP neural network and verify the feasibility and validity of the ensemble learning method for fault diagnosis. Meanwhile, the results also indicate that the proposed method can meet the requirements of a real-time diagnosis of the nuclear power plant.

1. Introduction

Structures of the nuclear power plant (NPP) are complicated, which have potential radiation detriment. Thus, the requirements for their safety and reliability are quite strict. One technical way to provide operation supports for the operators is fault diagnosis technology, and its application in NPP can assist operators to find and identify faults timely and accurately. It is also an effective method for preventing and reducing human factor errors. Fault diagnosis technology has important significance in improving the safe operation of NPP [1].

With the wide application of digital instruments and control systems in NPP, the fault diagnosis methods based on data driving have obtained extensive research. In the literature, a lot of fault diagnosis methods have been proposed; for example, Wang Hang applied the support vector machine (SVM) and improved particle swarm optimization (PSO) to perform further diagnosis in NPP on the basis of qualitative reasoning by knowledge-based preliminary diagnosis and sample data provided by an online simulation model [2]. Sinuhe adopted a time-lagged feed-forward

neural network in the research of the total instantaneous blockage of an assembly in the core of a sodium-cooled fast reactor [3]. Principal component analysis (PCA) is applied for fault detection of sensors in a nuclear power plant by Li et al. [4]. Qiu Pinda addressed a scheme for improving the accuracy of fault diagnosis by fusing the sparse coding fault diagnosis results with the expert system [5]. Shyamapada Mandal addressed an approach for small/minor fault detection of thermocouple sensors in a nuclear power plant using time series analysis methods [6]. He proposed another method based on the symbolic dynamic filter (SDF); the proposed method is also able to detect and classify the minor as well as major signal faults in the thermocouple signal [6]. To handle the complex sensor error detection problem in NPPs, an LSTM network was applied with successful results obtained by Choi and Lee [7].

The above methods have made some achievements in fault diagnosis of NPP, but they have their drawbacks: neural network based on gradient descend method is prone to get the local minimum with slow convergence and tendency to overfitting [8]; SVM lacks sparseness, and with the increasing sample size, the number of support vector tends to

increase with time spent in model training [9]. PCA is applied to fault detection on the basis of steady conditions, but for the variable conditions of NPP, a single PCA model is difficult to apply because it may cause misdiagnosis and missed diagnosis [10, 11]. Expert systems require a great deal of expert experience which is often hard to get [12]. SDF is essentially a kind of time series analysis method, and the basic idea of SDF is using temporal relation of historical data to predict behavior; therefore, SDF has better performances for strong time sequence data, but with limited capabilities for data with a lot of nonlinear relation. LSTM is difficult to determine parameters, so it is difficult to be practice applied [13].

Although a lot of methods have been done on fault diagnosis in NPP, most applications focus exclusively on the design and optimization of each method as well as the implementation of each method on various systems or components. That is, different algorithms have different performances, advantages, and disadvantages. It can be really difficult to satisfy all the needs in NPP with a single algorithm.

How to improve the performance of each algorithm has become the research hotspot of many scholars. Fortunately, the emergence of ensemble learning can provide ideas for solving those problems, but there is little research on the fault diagnosis with an ensemble learning algorithm in NPP. The model that meets the training samples may not have the same good performance in practical application, so the model will be exposed to a certain amount of risk when being output. It is really worth mentioning that the integration of multiple models can reduce this risk. Ensemble learning can combine multiple weak models to a good strong model. With the increasing of the number of weak models, the error rate of the integration decreases exponentially and eventually goes to zero [14, 15]. And ensemble learning can also greatly improve the stability and generalization ability of original weak models [16].

In this paper, as a representative algorithm of ensemble learning, the AdaBoost algorithm combined with the SAMME algorithm (with sagewise additive modeling using a multiclass exponential loss function) [17] is applied in fault diagnosis research in NPP, which selecting extreme learning machine (ELM) as the weak classifier. The contributions of this paper are as follows:

- (i) A fault diagnosis method based on ELM-AdaBoost.SAMME for the nuclear power plant is proposed.
- (ii) AdaBoost.SAMME is applied to improve the performance of the ELM algorithm.
- (iii) The novel part of this work is to verify the feasibility and validity of the ensemble learning method for fault diagnosis in this paper.

The paper is organized as follows: based on previous research, innovation points of this paper are described in Section 1. Two different methods are introduced to integrate into Section 2. Section 3 outlines the fault diagnosis framework in NPP with an ensemble learning algorithm.

The accuracy and effectiveness of the ensemble learning algorithm are evaluated in the HPR1000 simulator in Section 4. Conclusions are given in the last section.

2. Proposed Algorithm: ELM-AdaBoost.SAMME

2.1. ELM Algorithm. ELM algorithm is proposed based on SLFNs, which randomly selects the input weight matrix and hidden layer biases, and there is no need to adjust. This algorithm has many advantages; on one hand, the ELM algorithm can get the output weight of SLFNs through a simple generalized inverse operation of the hidden layer output matrices, which only needs to set the number of hidden layer neurons at the entire training process. On the other hand, the ELM algorithm is composed of simple mathematical algorithms; thus, it has a significant advantage in learning efficiency, and it overcomes the disadvantage of gradient-based traditional learning algorithm with falling into local minimum and overfitting [18]. The structure of ELM is shown in Figure 1.

Supposed that an original training data matrix with N variables can be expressed as $x_i = [x_{i1}, x_{i2}, \dots, x_{iN}]^T \in R_N$ and $y_i = [y_{i1}, \dots, y_{iK}]^T \in R_K$. For example, a ELM network with L hidden nodes and activation function $g(x)$ can be defined as

$$\sum_{i=1}^L \beta_i g(\omega_i \cdot x_j + b_i) = o_j, \quad i = 1, 2, \dots, L, \quad j = 1, 2, \dots, N, \quad (1)$$

where weights ω_i and bias b_i are randomly chosen and β_i is output weight which connects the i th hidden node with the output node; they are expressed in equation (2). Activation function $g(x)$ can select any bounded continuous functions like sigmoid, threshold, sine, and cosine:

$$\begin{aligned} \omega &= \begin{pmatrix} \omega_{11} & \dots & \omega_{1N} \\ \vdots & \ddots & \vdots \\ \omega_{L1} & \dots & \omega_{LN} \end{pmatrix}_{L \times N}, \\ \beta &= \begin{pmatrix} \omega_{11} & \dots & \omega_{1K} \\ \vdots & \ddots & \vdots \\ \omega_{L1} & \dots & \omega_{LK} \end{pmatrix}_{L \times K}, \\ b &= \begin{pmatrix} b_1 \\ \vdots \\ b_L \end{pmatrix}_{L \times 1}. \end{aligned} \quad (2)$$

The output of ELM can approximate the N samples with zero error, which can be described as

$$\sum_{i=1}^N \|o_i - y_i\| \approx 0. \quad (3)$$

Then, we can have

$$f_L(x) = \sum_{i=1}^L \beta_i g(\omega_i \cdot x_i + b_i) = y_i, \quad i = 1, 2, \dots, L. \quad (4)$$

Equation (4) can be expressed in a condensed form as $H\beta = Y$, where H is the hidden layer output matrix of the ELM:

$$H(\omega_1, \dots, \omega_L, b_1, \dots, b_L, x_1, \dots, x_N) = \begin{pmatrix} g(\omega_1 \cdot x_1 + b_1) & \dots & g(\omega_L \cdot x_N + b_L) \\ \vdots & \ddots & \vdots \\ g(\omega_1 \cdot x_N + b_1) & \dots & g(\omega_L \cdot x_N + b_L) \end{pmatrix}_{N \times L}. \quad (5)$$

ELM is aimed to minimize training error as well as the norm of output weights. Then, equation (5) is turned into a least-square problem: $\min_{\beta} \|H\beta - Y\|$. The calculation of output weights can be transformed into finding least-square solutions to the above linear equation, which is given as

$$\tilde{\beta} = H^+ Y, \quad (6)$$

where H^+ is the Moore–Penrose generalized inverse of matrix H .

Based on the foregoing analysis, the ELM algorithm has many specialties like the smallest norm of weights and the minimum approximation of training error and only needs to debug the number of hidden layers. All these advantages make it well suitable for the fault diagnosis of NPP in the following chapters.

2.2. Adaboost.SAMME Algorithm. Adaboost is a machine learning algorithm based on boosting [19]. Adaboost algorithm can integrate several classifiers to generate a new classifier with better performance. When solving the binary classification problem, the Adaboost algorithm only requires the accuracy of each weak classifier to be higher than 0.5 for each training sample. However, this requirement is much forced for multiclassification problems. Meanwhile, the weak classifiers are required to have slightly higher accuracy than random guesses; this requirement is also too weak which will not ensure the improving classification accuracy of the strong classifier. To solve multiclassification problems, SAMME is introduced into the Adaboost algorithm:

Firstly, each sample is given an initial weight value, then Adaboost.SAMME algorithm can adaptively update weight values by iteratively calculating the error rate of the weak classifier. The error rates also determine the weight value of each weak classifier, which represents its share in the final strong classifier. It can be briefly described as the fellow steps (see Algorithm 1).

Algorithm 1. Adaboost.SAMME algorithm.

Input: training samples (x_i, y_i) , $x_i \in R_N$, $y_i \in (1, 2, \dots, K)$, where K is the number of fault types, $G_m(x)$ is selected as the weak classifier based on the ELM method in this paper, and M is the number of weak classifiers.

(1) Initialize the weight value of the training samples:

$$\omega_{1i} = \frac{1}{N}, \quad i = 1, 2, 3, \dots, N. \quad (7)$$

(2) For $m = 1: M$

(1) Generate a weak classifier by training with weighted samples:

$$G_m(x): X \longrightarrow (1, 2, \dots, K), \quad (8)$$

$$\text{err}_m = \sum_{i=1}^N \|\omega_{mi} (G_m(x_i) \neq y_i)\|, \quad (9)$$

$$\alpha_m = \frac{1}{2} \log \frac{1 - \text{err}_m}{\text{err}_m} + \log(K - 1), \quad (10)$$

$$\omega_{m+1,i} = \frac{\omega_{mi}}{Z_m} \exp(-\alpha_m y_i G_m(x_i)), \quad i = 1, 2, \dots, N, \quad (11)$$

$$Z_m = \sum_{i=1}^N \omega_{mi} \exp(-\alpha_m y_i G_m(x_i)), \quad (12)$$

$$G(x) = \text{sign} \left(\sum_{m=1}^M \alpha_m G_m(x) \right). \quad (13)$$

(2) Calculate the error rate of each weak classifier:

(3) Calculate the weight value of each weak classifier:

(4) Renew the weight value of training samples according to step (3):

where Z_m is a standardization factor:

End.

(3) output: strong classifier prediction:

3. Design of Fault Diagnosis Method for NPP Based on ELM-AdaBoost.SAMME

3.1. Characteristic Parameter Selection for Fault Diagnosis. Taking HPR1000 NPP as an example, the research and experiments are based on the HPR1000 simulator. The typical faults of NPP selected for experimental verification are loss of coolant accident (LOCA), steam generator tube rupture (SGTR), main steam pipe rupture in the

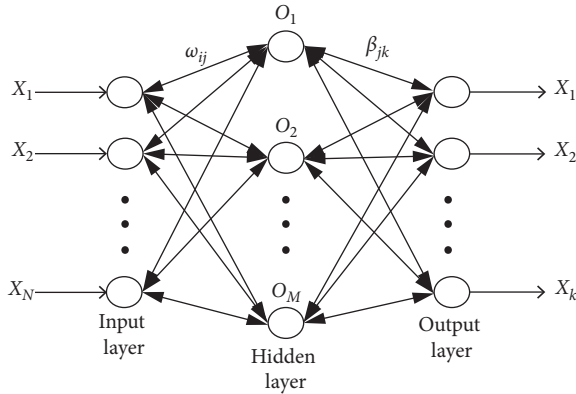


FIGURE 1: The structure of ELM.

containment, and main steam pipe rupture outside the containment. These faults are typical accidents that may destroy the integrity of NPP. If any accident occurs, NPP would undoubtedly be in an extremely dangerous state, and the operators may hardly judge the fault and take countermeasures quickly. The consequences would be unimaginable if mishandled. Through analyzing the reasons for the accidents happened, 21 fault symptoms are selected for and elaborated in the following paragraph:

The temperature inside the containment; containment sump level; pressurizer water level; no. 2 SG steam flow; coolant flow in loop 1; coolant flow in loop 3; no. 1 SG steam pressure; no. 3 SG steam pressure; temperature of the hot leg in loop 2; temperature of the cold leg in loop 1; temperature of the cold leg in loop 3; pressure inside the containment; pressurizer pressure; no. 1 SG steam flow; no. 3 SG steam flow; coolant flow in loop 2; primary average pressure; no. 2 SG steam pressure; temperature of the hot leg in loop 1; temperature of the hot leg in loop 3; and temperature of the cold leg in loop 2.

3.2. Framework for NPP Fault Diagnosis Based on ELM-AdaBoost.SAMME. The essence of fault diagnosis for NPP is a multiclassification problem, a combination of ELM and AdaBoost.SAMME can settle the problem well. ELM consists of three layers, the 21 characteristic parameters brought forward by the foregoing words are chosen as the sample inputs of ELM. Due to a large number of input parameters and training samples, the number of hidden layers is 140 which is decided by the simulation experiments in the following section. Four typical faults encoding are chosen as the output of ELM.

Algorithm flow based on the ELM-AdaBoost.SAMME is summarized in Figure 2. The left of the framework is the AdaBoost.SAMME algorithm; next to it is the ELM algorithm. During each pass in the loop, the AdaBoost.SAMME algorithm is called to generate a weak classifier applying ELM algorithm. After training the classifier with the whole samples, errors of the classifier are obtained and saved. The errors are used to calculate the weight values which will decide the classifiers' share in the final classifier and update samples' weight values. AdaBoost.SAMME algorithm will

stop automatically after the predetermined numbers are reached. Combining the whole classifiers, we can get the final output encoding as the result of fault diagnosis.

4. Simulation Tests and Analysis

In this section, the diagnosis model is trained and tested with experimental data to verify its ability of fault recognition. Meanwhile, in order to test the model performance on every aspect, BP neural network has been established for comparison.

4.1. Settings for ELM. A better activation function for ELM will make a contribution to the more accurate detection of a fault that occurred in NPP. From Figure 3, three activation functions are chosen. Sin function's error rate has been consistently high within the range of 100 hidden layers of ELM. Hardlim function is much the same with the sigmoid, but the curve of the error rate of the sigmoid function is more stable and with the smallest error rate which represents the highest accuracy, relative to other functions. Thus, the sigmoid function is selected as the activation function for the ELM algorithm.

The number of hidden neurons of the ELM algorithm is crucial to the performance and accuracy of the network, as well as restricting the training time. At the beginning of the training, the number of hidden neurons is 10, and then it gradually increases to 300 with a step of 10. Finally, the error rate curves of the training samples and the test samples with the different numbers of hidden neurons are shown in Figure 4, and the curve of training time is shown in Figure 5. From Figure 4, it can be seen that the error rate tends to decrease remarkably with the increasing number of hidden neurons, whether to the training curve or testing curve. Moreover, the accuracy of training samples is approximately keeping at 90%, when the number of hidden neurons reaches 100. The training time increases slightly with the number of hidden neurons and basically maintains at 1.76 s which is illustrated in Figure 5.

To verify the boosting effect of the ELM-AdaBoost.SAMME method, the following simulation experiments with the different numbers of weak classifiers are carried out. The three curves represent performances of ELM with the different numbers of hidden layers: 140, 230, and 400. 140 is set as the number of hidden layers for a weak classifier, 230 represents a relatively stronger weak classifier, and the strongest weak classifier has 400 hidden layers. The results under the same condition are illustrated in Figure 6 and Table 1. The result shows that the ELM with 140 hidden layers has the biggest jump with a 4.45% promotion of accuracy rate. The strongest weak classifier has the lowest error rate from the beginning; however, it has scant room for higher accuracy because of taking up too much computer resources. Thus, for the purpose of verifying algorithms' validity, there is no need to select the model with the smallest error rate. In a word, in order to demonstrate a more pronounced effect, the number of hidden neurons is selected as 140 with lower training accuracy of 89.25% in this paper.

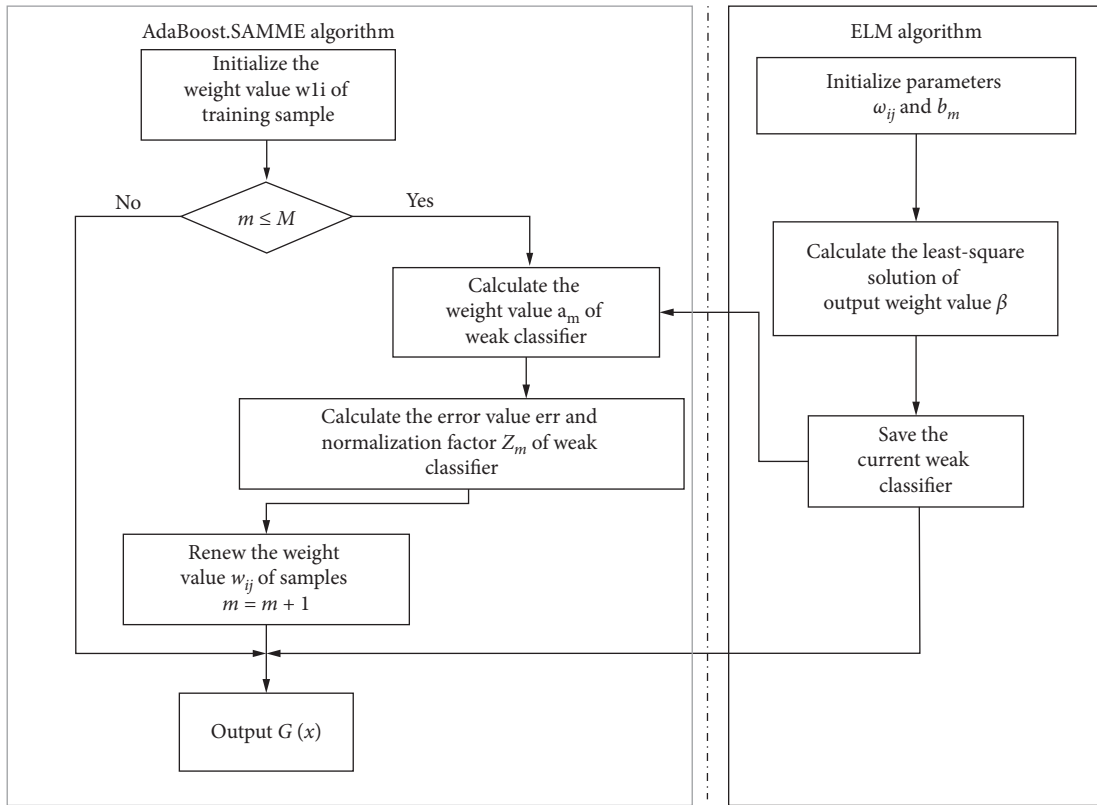


FIGURE 2: The framework of fault diagnosis based on ELM-AdaBoost.SAMME.

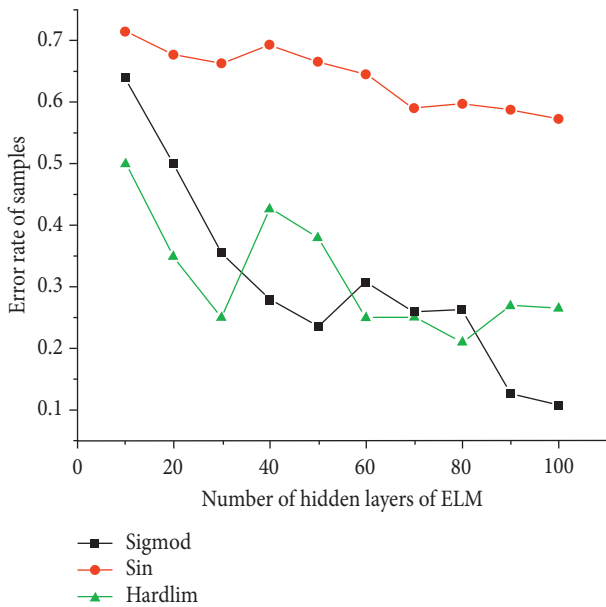


FIGURE 3: Error rate of samples with different activation functions of ELM.

4.2. *Settings for ELM-AdaBoost.SAMME.* The relationship between the error rate and the number of weak classifiers is shown in Figure 7. The error rate is the ratio of the number of wrong classified samples and all samples.

According to Figure 7, we can see that the error rate of training samples tends to decrease with the increasing

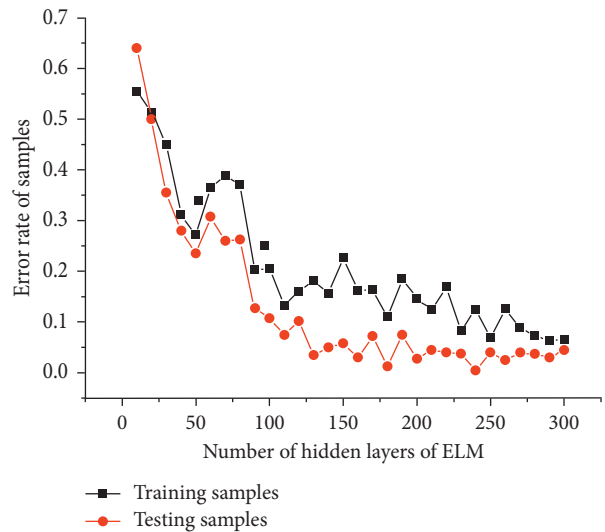


FIGURE 4: Error rate of samples with a different number of hidden layers of ELM.

number of weak classifiers, which demonstrates the effect of the AdaBoost.SAMME algorithm to bring the error down. It is also clearly observed that the error rate of testing samples is firstly remarkably decreased and then increased slowly, which indicates that error rate reaches the lowest value with the increasing of the number of weak classifiers during the initial phases; however, followed by the continuous increasing of weak classifiers, the error rate of testing samples

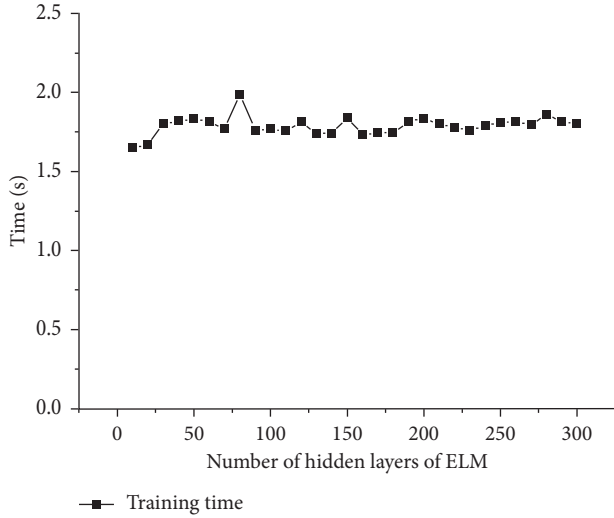


FIGURE 5: Training time of samples with a different number of hidden layers of ELM.

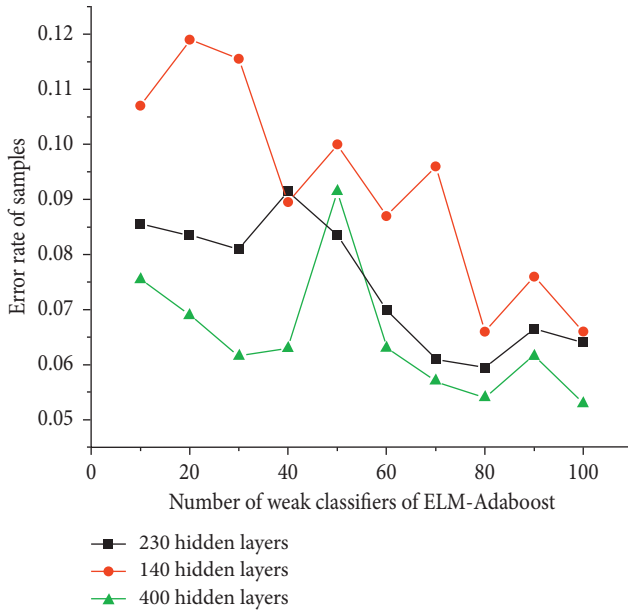


FIGURE 6: Error rates of samples with a different number of hidden layers and weak classifiers.

increases because of model overfitting. Meanwhile, due to the increase in the number of weak classifiers, the training time also increases. Based on this analysis, the number of weak classifiers for the ELM-AdaBoost.SAMME algorithm can be set as 100 with the lowest testing error rate since it also regards the curves of error rate (Figure 7) and training time (Figure 8) as selection criteria.

4.3. Comparative Analysis of Simulation Results

4.3.1. Model Comparison and Analysis. This section presents the results of the proposed method, the efficiency of the proposed method is validated by the data of the HPR1000 simulator using the indigenous design

simulation platform of China. The period of data capture is set as 4 min and then 500 samples are used to train the ELM-AdaBoost.SAMME model and randomly select 200 samples for testing.

Training samples are made up of two parts: the input part is fault symptom variables and the output part is fault indication. Training samples include the four-fault data as well as the normal operating state. All the output samples of fault indication are encoded with one-hot code. For example, the encoding [1 0 0 0] means LOCA has occurred. 1 stands for fault happening and 0 otherwise.

For comparing the properties of the ELM-AdaBoost.SAMME model with other algorithms, the simulation tests are carried out under the MATLAB environment on a PC platform with 2 Intel Core dual-core CPUs (3.6 GHz) and 16 GB RAM, and the results of BP neural network and ELM with the same test conditions are illustrated in Table 2. Especially for validating the advantages of the combined algorithm ELM and AdaBoost.SAMME, the BP-AdaBoost.SAMME is selected for characteristic comparison.

The calculation formulas of the accuracy rate are presented as follows:

$$e = \frac{(N_{\text{all}} - N_{\text{err}})}{N_{\text{all}}}, \quad (14)$$

where N_{all} is the total number of samples and N_{err} is the number of wrong classified samples.

As shown in Table 2, it can be concluded as follows:

- (1) Simulation results show that ELM-AdaBoost.SAMME model achieves 97.5% precision in the testing stage, which is better than other methods. In contrast with ELM, the accuracy of ELM-AdaBoost.SAMME method, in training and testing stage, rises by 4.25% and 5%, respectively, at the same time, and there is no significant increase in time.
- (2) BP neural network has the highest accuracy in the training stage but spends the maximum time. In addition, the problem of overfitting appears during the testing stage for the BP neural network.
- (3) The ELM algorithm has the feature of training quickly because it can be done without iterative tuning, which has dozens of times faster than the BP neural network. What is more, the accuracy of the ELM algorithm can achieve a high degree through training.
- (4) By comparison, the accuracy of ELM-AdaBoost.SAMME and BP-AdaBoost.SAMME is improved in different degrees separately on training and testing set; however, the BP-AdaBoost.SAMME has lower accuracy on the testing set. It could be inferred that the BP neural network had been overfitted; moreover, the training time has almost doubled, to 2 minutes.

Then, the conclusion can be drawn that the AdaBoost.SAMME combined with the ELM algorithm as its weak classifiers can be adopted to achieve greater speed, as well as accuracy.

TABLE 1: Simulation results of different numbers of weak classifiers and hidden layers.

Number of hidden layers	Accuracy with 10 weak classifiers	Accuracy with 100 weak classifiers	Promotion of accuracy rate (%)
140	0.8905	0.935	4.45
230	0.9145	0.936	2.15
400	0.9245	0.947	2.25

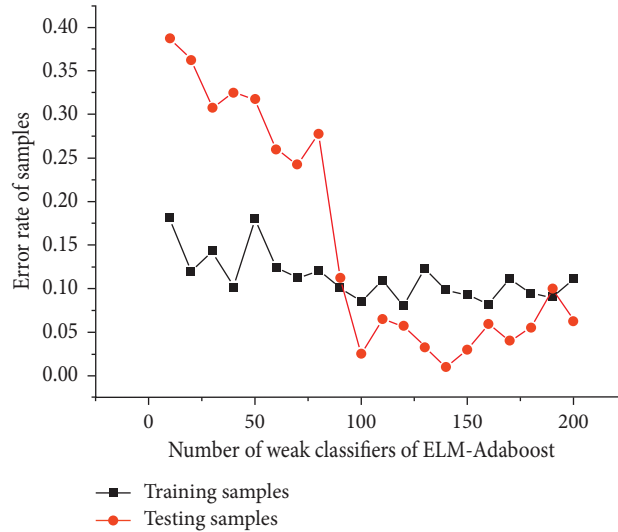


FIGURE 7: Error rate of samples with a different number of weak classifiers of ELM-AdaBoost.SAMME.

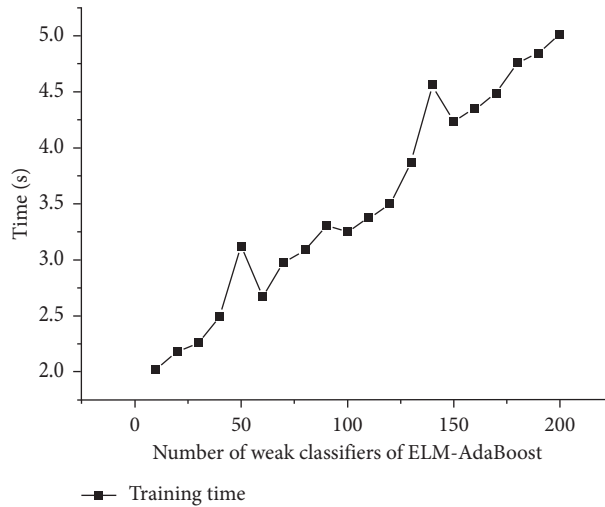


FIGURE 8: Training time of samples with a different number of weak classifiers of ELM-AdaBoost.SAMME.

4.3.2. *Online Simulation Experiment.* In this section, in order to test the real-time performance of the ELM-AdaBoost.SAMME model, the ELM-AdaBoost.SAMME and BP neural network models based on C programming language are both integrated into the HPR1000 simulator. Online simulation tests are carried out in the HPR1000 simulator under normal operating conditions with full power, and then LOCA malfunction is inserted by instructor station (IS) as an example after the simulator running for 20 s. Based on the simulation results in Figure 9, the red curve yout1 (2) represents the output variable of the diagnosis model based

on the ELM-AdaBoost.SAMME algorithm indicating the appearance of LOCA, whereas the orange curve yout2 (2) represents the output variable based on the BP neural network. The other colored curves are for the rest of the fault diagnosis variables.

As we can see clearly in Figure 9, the ELM-AdaBoost.SAMME model diagnoses the malfunction of LOCA in 25 s after inserting malfunction in 20 s, and the BP model diagnoses the malfunction in 28 s. Both models can diagnose the malfunction correctly; nonetheless, the ELM-AdaBoost.SAMME is 3 seconds faster than the BP model.

TABLE 2: Simulation results of three algorithms.

Algorithm	Average accuracy of training samples	Average accuracy of testing samples	Average training time/s	Model parameters
ELM	0.8925	0.925	1.6018	Hidden layer number: 140
ELM-AdaBoost.SAMME	0.935	0.975	3.3401	Weak classifier number: 100 Hidden layer number: 140
BP	0.9638	0.918	46.7853	Hidden layer number: 45
BP-AdaBoost.SAMME	0.977	0.921	102.9676	Hidden layer number: 45 Hidden layer number: 100

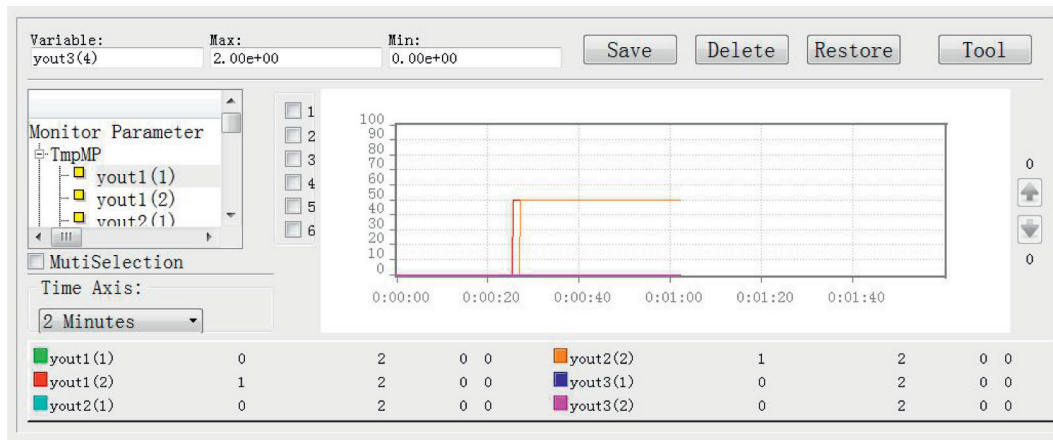


FIGURE 9: Comparison curves of an online test of the two models.

With the foregoing analysis, we can conclude that the ELM-AdaBoost.SAMME model can diagnose the malfunction more quickly and reliably. Therefore, the AdaBoost.SAMME algorithm really does have some accuracy enhancements for weak classifiers. At the same time, it is proved that the ELM-AdaBoost.SAMME model can also meet the real-time requirement.

5. Conclusions

In this paper, an ELM-AdaBoost.SAMME-based method is applied for fault diagnosis in NPP. The ELM-AdaBoost.SAMME model selects the ELM algorithm as its weak classifiers and then applies the AdaBoost.SAMME algorithm to integrate weak classifiers into a strong classifier. To estimate the fault diagnosis ability of the proposed model for NPP, a diagnosis model based on ELM-AdaBoost.SAMME is built. What is more, the model integrates into the HPR1000 simulator for real-time testing and a more accurate reflection. Simulation results show that the model is characterized by the fast and accurate diagnosis as well as real-time requirement. Meanwhile, the AdaBoost.SAMME algorithm is also proved to be effective in boosting the weak classifier. That is, the

AdaBoost.SAMME algorithm can be applied not only in the ELM algorithm but also in other diagnosis algorithms.

Data Availability

The data for this paper came from the simulator.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Y. Yao, J. Wang, M. Xie, L. Hu, and J. Wang, "A new approach for fault diagnosis with full-scope simulator based on state information imaging in nuclear power plant," *Annals of Nuclear Energy*, vol. 141, 2020.
- [2] H. Wang, M.-J. Peng, J. Wesley Hines, G.-Y. Zheng, Y.-K. Liu, and R. Upadhyaya Belle, "A hybrid fault diagnosis methodology with support vector machine and improved particle swarm optimization for nuclear power plants," *ISA Transactions*, vol. 95, 2019.
- [3] S. Martinez-Martinez, N. Messai, J.-P. Jeannot, and D. Nuzillard, "Two neural network based strategies for the detection of a total instantaneous blockage of a sodium-cooled

- fast reactor,” *Reliability Engineering and System Safety*, vol. 137, 2015.
- [4] W. Li, M.-J. Peng, and Q.-Z. Wang, “False alarm reducing in PCA method for sensor fault detection in a nuclear power plant,” *Annals of Nuclear Energy*, vol. 118, 2018.
- [5] P. Qiu and L. Ye, “Research on fault diagnosis method of coolant system in nuclear power plant,” in *Proceedings of 2019 3rd International Conference on Artificial Intelligence, Systems, and Computing Technology (AISCT 2019)*, pp. 477–481, Institute of Management Science and Industrial Engineering, Brasilia, Brazil, October 2019.
- [6] S. Mandal, B. Santhi, S. Sridhar, K. Vinolia, and P. Swaminathanm, “Minor fault detection of thermocouple sensor in nuclear power plants using time series analysis,” *Annals of Nuclear Energy*, vol. 134, 2019.
- [7] J. Choi and S. J. Lee, “Consistency index-based sensor fault detection system for nuclear power plant emergency situations using an LSTM network,” *Sensors*, vol. 20, no. 6, 2020.
- [8] W. Feng-qin, Y. Jian-guo, S. Xiao-li, and L. Yu, “Application of an improved BP neural network in fault diagnosis of aircraft fuel system,” *Computer Measurement and Control*, vol. 14, no. 7, pp. 867–871, 2006.
- [9] M. E. Tipping, “Sparse Bayesian learning and the relevance vector machine,” *Journal of Machine Learning Research*, vol. 1, no. 3, pp. 211–244, 2001.
- [10] J.-C. Jeng, C.-C. Li, and H.-P. Huang, “Fault detection and isolation for dynamic processes using recursive principal component analysis (PCA) based on filtering of signals,” *Asia-Pacific Journal of Chemical Engineering*, vol. 2, no. 6, pp. 501–509, 2007.
- [11] M. Navi, N. Meskin, and M. Davoodi, “Sensor fault detection and isolation of an industrial gas turbine using partial adaptive KPCA,” *Journal of Process Control*, vol. 64, no. 21, pp. 37–48, 2018.
- [12] J. Burriel-Valencia, R. Puche-Panadero, and J. Martinez-Roman, “Automatic fault diagnostic system for induction motors under transient regime optimized with expert systems,” *Electronics*, vol. 8, no. 1, 2018.
- [13] M. Tan, S. Yuan, S. Li, Y. Su, H. Li, and F. H. He, “Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning,” *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 2937–2948, 2020.
- [14] G.-kai Ma, WuLi-gang, and Y. Wang, “A general subspace ensemble learning framework via totally-corrective boosting and tensor-based and local patch-based extensions for gait recognition,” *Pattern Recognition*, vol. 66, 2017.
- [15] C.-W. Deng, G.-B. Huang, J. Xu, and J.-X. Tang, “Extreme learning machines: new trends and applications,” *Science China (Information Sciences)*, vol. 58, no. 2, pp. 5–20, 2015.
- [16] Li Yang and Z. Yang, “Application of EOS-ELM with binary jaya-based feature selection to real-time transient stability assessment using PMU data,” *IEEE Access*, vol. 5, pp. 23092–23101, 2017.
- [17] A. Fernández-Baldera, J. M. Buenaposada, and B. Luis Baumela, “AdaCost: multi-class boosting with costs,” *Pattern Recognition*, vol. 79, 2018.
- [18] G. B. Huang, Q. Y. Zhu, and C. K. Siew, “Extreme learning machine: a new learning scheme of feedforward neural networks,” in *Proceedings of 2004 IEEE International Joint Conference on Neural Networks*, pp. 985–990, IEEE, Budapest, Hungary, July 2004.
- [19] P. Zhang and J. Yan, “Fault feature extraction of wind turbine rolling bearing based on PSO-VMD,” in *Proceedings of the 3rd International Conference on Electrical, Automation and Mechanical Engineering*, Guizhou, China, June 2019.