Article

Fault Diagnosis of Nuclear Power Plant Based on Sparrow Search Algorithm Optimized CNN-LSTM Neural Network

Chunyuan Zhang [1,2], Pengyu Chen [1,2], Fangling Jiang [3], Jinsen Xie [1,2,*] and Tao Yu [1,2,*]

1   School of Nuclear Science and Technology, University of South China, Hengyang 421000, China
2   Research Center for Digital Nuclear Reactor Engineering and Technology of Hunan Province, University of South China, Hengyang 421000, China
3   College of Computer Science, University of South China, Hengyang 421000, China
*   Correspondence: jinsen_xie@usc.edu.cn (J.X.); yutao29@sina.com (T.Y.)

Abstract: Nuclear power is a type of clean and green energy; however, there is a risk of radioactive material leakage when accidents occur. When radioactive material leaks from nuclear power plants, it has a great impact on the environment and personnel safety. In order to enhance the safety of nuclear power plants and support the operator's decisions under accidental circumstances, this paper proposes a fault diagnosis method for nuclear power plants based on the sparrow search algorithm (SSA) optimized by the CNN-LSTM network. Firstly, the convolutional neural network (CNN) was used to extract features from the data before they were then combined with the long short-term memory (LSTM) neural network to process time series data and form a CNN-LSTM model. Some of the parameters in the LSTM neural network need to be manually tuned based on experience, and the settings of these parameters have a great impact on the overall model results. Therefore, this paper selected the sparrow search algorithm with a strong search capability and fast convergence to automatically search for the hand-tuned parameters in the CNN-LSTM model, and finally obtain the SSA-CNN-LSTM model. This model can classify the types of accidents that occur in nuclear power plants to reduce the nuclear safety hazards caused by human error. The experimental data are from a personal computer transient analyzer (PCTRAN). The results show that the classification accuracy of the SSA-CNN-LSTM model for the nuclear power plant fault classification problem is as high as 98.24%, which is 4.80% and 3.14% higher compared with the LSTM neural network and CNN-LSTM model, respectively. The superiority of the sparrow search algorithm for optimizing model parameters and the feasibility and accuracy of the SSA-CNN-LSTM model for nuclear power plant fault diagnosis were verified.

Keywords: convolutional neural network; fault diagnosis; long short-term memory; nuclear power plant; sparrow search algorithm

1. Introduction

Clean and low carbon energy is the trend of global energy development and the internal requirement of high-quality energy development. In the context of carbon neutrality, the development of energy will undergo profound changes. As a green and clean energy source, nuclear energy has the advantages of high energy density, non-intermittency, and less constraints from natural conditions. In particular, it is important to reduce the use of traditional fossil energy, reduce carbon emissions, and achieve carbon neutrality [1].

While vigorously developing nuclear energy, ensuring nuclear safety is the primary prerequisite [2]. For nuclear power systems, the energy source is the nuclear reactor, in which nuclear fission reaction occurs. Compared with traditional thermal power plants, the particularity of nuclear power plants is not only reflected in the complexity of nuclear power systems, but also in the risk of radioactive material leakage after its accidents [3,4]. Therefore, a reliable shielding design is particularly important [5]. In the event of an

accident in a nuclear power plant, operators are required to make timely and accurate judgments. Since the complexity of nuclear power plants and unpredictability of accidents are beyond the basis for design, it is very important to design an optimization system with high reliability and strong applicability [6]. It is of great practical significance to use auxiliary decision techniques to support the operator's activity under the accident condition, reduce the possibility of the operator making false judgments, and to avoid severe accidents.

With society's increasing attention to the nuclear power industry, people's requirements for nuclear safety are also increasing. Fault diagnosis technology plays an important role in the field of nuclear power plant fault diagnosis. Fault diagnosis methods can be broadly divided into quantitative and qualitative analysis methods, among which common quantitative analysis methods include neural networks, support vector machines, rough set theory, etc. The neural network-based fault diagnosis method can handle nonlinear problems, has parallel computing capabilities, and does not require diagnostic and inference rules; it can learn by mapping between the input and output of samples to obtain a training model. The support vector machine method can achieve better performance with less data and avoid overtraining. Cao et al. combined a support vector machine (SVM) with principal component analysis (PCA) to establish a three-layer fault classification model, and simulated three faults in small pressurized water reactors (SPWRs). The experimental results showed that the method had rapidity and high accuracy [7]. Zio et al. proposed a fault classification method for nuclear power plants based on support vector machines (SVMs). This method combined single-class support vector machines and multi-class support vector machines into a hierarchical structure to classify boiling water reactor feedwater system faults [8]. Rough set does not need additional information and prior knowledge. Mu et al. proposed a method based on neighborhood rough set to learn and diagnose the training sets of typical faults in nuclear power plants. The results showed that the method can quickly and accurately diagnose fault types [9]. Xu et al. proposed a fault diagnosis method for nuclear power plants based on support vector machines (SVMs) and rough sets (RSs). This method used rough sets to simplify the data, and then used support vector machines for fault classification [10].

Qualitative analysis methods include expert system, fuzzy logic, etc. Zhang et al. solved the nuclear power plant fault diagnosis problem with a frequency-based on-line expert system (FBOLES) that accurately detected abnormal signals in all 33 faults simulated [11]. Mwangi et al. discussed the adaptive neuro-fuzzy inference system (ANFIS) based on the fuzzy logic method. The small-break loss-of-coolant accident of Qinshan I Nuclear Power Plant was used for modeling, and the model had good prediction ability and high sensitivity [12]. The advantages and disadvantages of various methods are listed in Table 1.

In order to improve the safety of nuclear power plants, assist operators in fault identification and fault analysis, help operators to make corresponding operations more quickly and accurately in the event of an accident, and to avoid more serious accidents, we need to continuously develop and update the fault diagnosis technology of nuclear power plants. This paper innovatively proposes the SSA-CNN-LSTM model to solve the fault diagnosis problem of nuclear power plants.

Many fault diagnosis methods have been developed, but improving the accuracy of the model, finding more suitable optimization algorithms, and improving the generalization ability of the model are still the main directions of the research on fault diagnosis methods. In recent years, there have been many researches on metaheuristic algorithms. The metaheuristic algorithm is an algorithm based on intuitive or empirical construction, which can give a feasible solution to the proposed problem. The meta-heuristic strategy adopted by meta-heuristic optimization algorithm is usually a general heuristic strategy, which can be widely used in various fields. Many metaheuristic algorithms are inspired by phenomena in nature. The literature [13] details nine categories of metaheuristic algo-

rithms and they are biology-based, swarm-based, sports-based, music-based, social-based, math-based, physics-based, chemistry-based, and hybrid methods, as shown in Figure 1.

Table 1. Comparison of various methods.

|	Methods	Advantage	Shortcoming
Qualitative analysis approach	Fault tree	It can make a comprehensive and concise description of the various causes and logical relationships that lead to accidents.	For complex systems with more branching fault trees, the calculation is more complex, which brings some difficulties to qualitative and quantitative analysis.
	Expert System	It does not rely on system models, does not rely on large amounts of data, and can reason on the basis of insufficient information.	It requires a lot of knowledge and experience, and cannot diagnose unknown faults.
	Fuzzy Logic	It is suitable for the diagnosis of fuzzy phenomena and uncertain information systems, and can accurately diagnose failure modes that are closer to the fuzzy rules.	It is not good for the newly emerged fault diagnosis and does not have the ability to learn, and the rules of fuzzy relations are more difficult to establish.
Quantitative analysis method	Neural Networks	It has strong nonlinear mapping capability and parallel processing capability, which can solve the fault diagnosis problem of complex nonlinear systems.	It requires a large number of parameters, and the selection of training samples affects the performance of the network and cannot explain its own inference process.
	Support vector machines	It has a strong theoretical foundation, can avoid overtraining, can achieve better performance with less data, and can extract the maximum amount of classification knowledge from the data with limited feature information.	It cannot explain its own reasoning process, has limited types of solvable problems, and has a narrow field of application.
	Rough Logic	It does not require additional information and a priori knowledge.	It is not suitable for handling noisy data and has a weak fault tolerance.
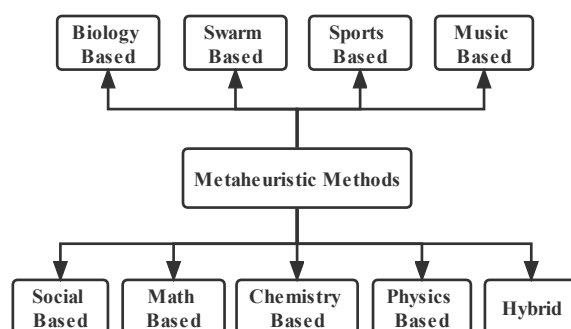
Figure 1. Metaheuristic algorithm classification.

People have constructed corresponding biology-based optimization algorithms based on biological behaviors or phenomena, such as invasive weed optimization (IWO) [14], photosynthetic algorithm [15], etc. Swarm-based intelligent optimization algorithms that have been built based on the group behavior of animals in nature include the firefly algorithm [16,17], bat algorithm [18], gray wolf optimization algorithm [19], ant lion optimization algorithm [20], whale optimization algorithm [21], etc. In accordance with

the physical principles or phenomena, physics-based optimization algorithms have been constructed, such as the simulated annealing algorithm [22], black hole algorithm [23], etc. There are also others such as the social-based imperialist competitive algorithm [24], music-based harmony search [25], chemistry-based chemical-reaction-inspired metaheuristic for optimization [26], etc.

Various metaheuristic algorithms are widely used in various fields, such as in computers, industry, medicine, economics, biology, etc. Optimization algorithms are also widely used in the field of nuclear engineering to solve problems. Amm et al. used the ant colony optimization algorithm to solve the nuclear core fuel reload optimization problem [27]. Khoshahval F et al. studied the application of particle swarm optimization and genetic algorithm in the nuclear fuel loading pattern problem and proved the effectiveness of the algorithm [28], etc.

In this paper, we combined the advantages of the convolutional neural network (suitable for processing complex data) and the long short-term memory neural network (suitable for processing time series data) to form the CNN-LSTM model, which was used to solve the problem of nuclear power plant fault diagnosis, with a classification accuracy that could reach 95.16%. We then used the sparrow search algorithm to optimize some parameters in the CNN-LSTM model to obtain the SSA-CNN-LSTM model. The experimental results show that the model which was optimized by the sparrow search algorithm is better than the CNN-LSTM model in performance, and proves the accuracy and feasibility of the SSA-CNN-LSTM model.

The fault diagnosis method proposed in this paper can help operators reduce human errors in the event of an accident and reduce the pressure on operators resulting from the accidents, which is of great significance for improving the safety of nuclear power plants. Compared with the traditional fault diagnosis method, the fault diagnosis method based on machine learning has the advantages of fast processing of large amounts of data, analysis and extraction of effective information, and good stability. As a result, in the fault diagnosis technology, the fault diagnosis technology based on machine learning method has received more and more attention. Compared with traditional machine learning methods, SSA is added to automatically optimize some parameters of the model to obtain the optimal model; this can improve the classification accuracy of the model. In addition, the model with SSA has the characteristics of fast convergence.

The rest of this paper is organized as follows. Section 2 introduces the basic principles of the CNN, the LSTM neural network and the sparrow search algorithm, and the construction of the CNN-LSTM and SSA-CNN-LSTM models. Section 3 presents the experimental data, the experimental analysis, and the experimental results. Section 4 presents the conclusion.

2. Methodology

2.1. Convolutional Neural Network (CNN)

The CNN is a feedforward neural network including convolution computation. The core of the CNN is the convolution kernel. The convolution layer uses the convolution kernel to extract data characteristics. Each neuron in each layer of the CNN characteristic mapping is only related to a small part of the neurons in the previous layer. In the convolution layer operation, the CNN greatly reduces the number of parameters and improves the model training speed by means of local connection of neurons and convolution kernel weight sharing. The structure of the CNN is usually composed of the convolution layer, pooling layer, and full connection layer. The convolution layer is composed of several characteristic graphs obtained by a convolution operation, as shown in Figure 2. The formula of the convolution layer is shown in Formula (1):

$$C_j^l = \sigma\left(\sum_{i \in M} x_i^{l-1} \times W_{ij}^l + b_j^l\right) \tag{1}$$

where l represents the number of layers. C_j^l represents the jth neuron of layer l. M represents the number of neurons connected between the previous layer and the current layer. W_{ij}^l is the weight. b_j^l is the bias.
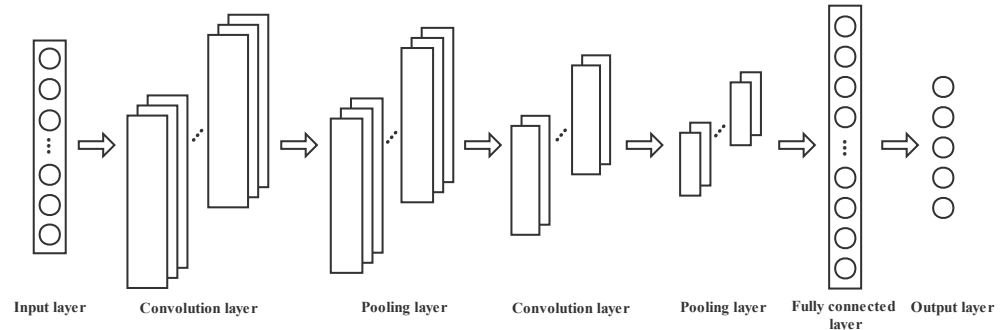

Figure 2. The network structure of the CNN.

The pooling layer is behind the convolution layer, and is used to compress the model, thus improving the robustness and calculation speed of the model and preventing the occurrence of overfitting to a certain extent.

2.2. Long Short-Term Memory (LSTM) Neural Network

The LSTM neural network introduces the concept of gating units based on the traditional recurrent neural network. When time series data is transferred between units in the implicit layer, it controls the degree of memory and forgetting of the previous and current data in the time series data through controllable gates such as forgetting gates, input gate, and output gate, so that the neural network has the function of long short-term memory. The LSTM neural network has good analysis ability for time series data, and it effectively improves the gradient disappearance and gradient explosion of recurrent neural networks.

The forget gate (f_t) is calculated by the hidden state of the last moment (h_{t-1}) and the input value of the current moment (x_t) through the sigmoid activation function layer. The hidden state at the last moment (h_{t-1}) and the input value at the current moment (x_t) get the input gate (i_t) through the sigmoid activation function layer, and the candidate cell state ($\hat{c}_t$) through the tanh activation function layer. The current cell state (c_t) is calculated from the last cell state (c_{t-1}), the new cell state ($\hat{c}_t$), and the input gate (i_t). The output gate (o_t) is calculated from the hidden state of the last moment (h_{t-1}) and the input value of the current moment (x_t). The output gate (o_t) and the current cell state (c_t) are calculated to obtain the current hidden state (h_t), as shown in Figure 3.
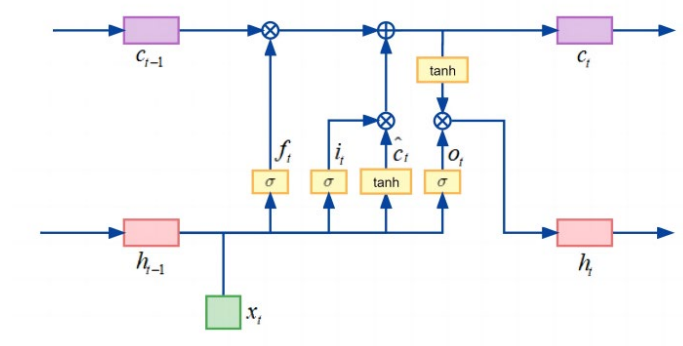

Figure 3. LSTM structure diagram.

The input gate is used to control the extent to which the current calculation state is updated to the memory cell. The input gate is calculated as shown in Equations (2) and (3).

$$\hat{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{2}$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{3}$$

where x_t is the input, h_{t-1} is the hidden state, W_c, W_i, U_c, U_i are the weight matrix and b_c, b_i are the bias.

The forget gate is used to control the extent to which the state of the input and current calculation is updated to the memory unit. The forgetting gate calculation formula is shown in (4):

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{4}$$

where x_t is the input, h_{t-1} is the hidden state, W_f, U_f are the weight matrix, and b_f is the bias.

The state is deleted and updated by the forget gate and the input gate. The state calculation formula is shown in (5):

$$c_t = f_t c_{t-1} + i_t \hat{c}_t \tag{5}$$

The output gate is used to control the input and the current output depending on the degree of the current memory unit. The output gate calculation formula is shown in (6) and (7):

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{6}$$

$$h_t = o_t \cdot \tan c_t \tag{7}$$

where x_t is the input, h_{t-1} is the hidden state, W_o, U_o are the weight matrix, and b_o is the bias.

2.3. CNN-LSTM Model

LSTM neural networks solve the problem of gradient explosion and gradient disappearance in traditional recurrent neural networks. Usually, a LSTM neural network is used to process time series data and predict classification. This paper attempts to use a LSTM neural network to learn the time series data of nuclear power plants and classify the types of nuclear power plant accidents.

Due to the complexity and diversity of nuclear power plant operation data, complex data will increase the training time of the neural network and affect the final classification results. However, a CNN is suitable for processing a large number of high-dimensional and non-linear data [29]. After a series of dimensionality reduction and feature extraction processing such as the CNN layer, batch normalization layer, activation function layer, and pooling layer, the operation data of nuclear power plants not only greatly reduce the number of parameters and retain the important features in the original data, but also improve the learning efficiency of the subsequent LSTM neural network and the fault diagnosis accuracy of the overall model. Therefore, this paper proposes the addition of a convolutional neural network (CNN) before the LSTM neural network to reduce the dimension of the data, thereby forming a CNN-LSTM model. The structure of the CNN-LSTM model is shown in Figure 4. The data input goes through the convolution layer, batch normalization layer, activation function layer, pooling layer, LSTM layer, fully connected layer, softmax layer, in turn, and finally outputs the results.

2.4. SSA-CNN-LSTM Model

In the above, the CNN-LSTM model was constructed to deal with the fault classification problem of nuclear power plants. However, in the process of constructing the model, it was found that there were many hyperparameters in the neural network, such as the number of neurons, learning rate, number of neural network layers, epoch, etc. The learning rate and the number of hidden layer neurons need to be manually debugged according to experience, and the setting of these two parameters will also have a greater impact on the accuracy of the final results. Therefore, it is of great significance to select an optimization algorithm that can automatically optimize the hyperparameters to improve the accuracy of model classification.
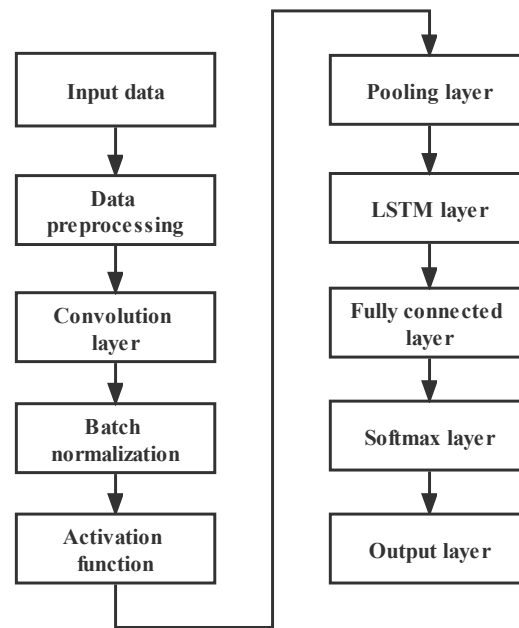
Figure 4. Flow Chart of the CNN-LSTM Model.

The sparrow search algorithm is a group intelligence optimization technique based on sparrow foraging and anti-predation behaviors [30]. It has a good search ability in solving optimization problems and has strong parallelism and stability. Therefore, SSA is selected to optimize the parameters of the CNN-LSTM model.

The behavior of sparrows is idealized and corresponding rules are formulated. The foraging process of sparrows can be abstracted as the producer and the scrounger. The producers are responsible for finding food in the population and providing a foraging area and direction for the entire sparrow population, while the scroungers receive food from the producers.

The location update of the producer can be expressed as shown in Formula (8):

$$Y_{i,j}^{t+1} = \begin{cases} Y_{i,j}^t \cdot \exp(-\frac{i}{\alpha \cdot iter_{\max}}) & if & R_2 < ST \\ Y_{i,j}^t + P \cdot L & if & R_2 \geq ST \end{cases} \tag{8}$$

where t is the current epoch, $j = 1, 2, 3, \ldots, d$. $iter_{\max}$ is the max number of epochs. $Y_{i,j}$ represents the location information in the jth dimension of the ith scrounger. $R_2 \in [0, 1]$ and $ST \in [0.5, 1]$ represent the early warning value and safety value, respectively. $\alpha \in (0, 1]$ is a random number. P is a random number subject to normal distribution. L is a matrix of all 1.

When $R_2 < ST$, this means there is no danger around at this time, and the scroungers can search in a wide range. If $R_2 \geq ST$, this indicates that there is a danger at this time and an alarm is issued. At this time, the population is transferred to a safe place.

The location update of the producer can be expressed as shown in Formula (9):

$$Y_{i,j}^{t+1} \begin{cases} Q \cdot \exp(\frac{Y_{worst} - Y_{i,j}^t}{i^2}) & if & i > n/2 \\ Y_p^{t+1} + \left| Y_{i,j}^t - Y_p^{t+1} \right| \cdot A^+ \cdot L & otherwise \end{cases} \tag{9}$$

where Y_p is the optimal position occupied by the producer, and Y_{worst} is the worst position. A denotes a matrix of $1 \times d$. Each element is 1 or -1 and $A^+ = A^T (AA^T)^{-1}$. When $i > n/2$, this indicates that the ith scrounger with a low fitness value did not receive food and was in a very hungry state. At this time, the producer needs to find new places to feed.

In case of danger, some sparrows will display anti predatory behavior, and its location update can be expressed as shown in Formula (10):

$$Y_{i,j}^{t+1} = \begin{cases} Y_{best}^{t} + \beta \cdot \left| Y_{i,j}^{t} - Y_{best}^{t} \right| & if \quad f_i > f_g \\ Y_{i,j}^{t} + K \cdot \left(\frac{\left| Y_{i,j}^{t} - Y_{worst}^{t} \right|}{(f_i - f_w) + \varepsilon} \right) & if \quad f_i = f_g \end{cases} \tag{10}$$

where Y_{best} is the current global optimal position. As a step control parameter, β is a random number subject to the normal distribution with mean value of 0 and variance of 1. $K \in [-1, 1]$ is a random number. f_i is the fitness value of the current sparrow individual. f_w and f_g are the global worst and best fitness values, respectively. ε is the smallest constant.

When $f_i > f_g$, this indicates that the sparrows are vulnerable to predator attacks. Y_{best} indicates that this position is the best position. When $f_i = f_g$, this shows that the sparrow is in danger. K indicates the direction of the sparrow's movement.

The relevant rules of the sparrow search algorithm are as follows:

(1) The producers are responsible for the task of searching for objects and paths.
(2) When the sparrow finds a predator, it will send an alarm to the population. When the alarm value is less than that of the safe value, the signal will be ignored, and when the alarm value is beyond that of the safe value, the population will go into play to escape to a safe area.
(3) The identities of the producers and scroungers of the sparrow population are not fixed, but the proportion of the discoverers is fixed.
(4) The lower the fitness value of the scroungers in the population, the worse their position will be in the population, indicating that they will need to forage elsewhere.
(5) Scroungers can find producers that provide better foraging areas in the sparrow population.
(6) When the population is threatened, individuals at the edge will find a safe position and move, while individuals at other positions in the population move randomly.

The fitness function is an important part of the optimization problem, which can measure the performance of the algorithm. The sparrow search algorithm calculates the fitness value once at each population update to determine the classification accuracy of the sparrow search algorithm. The fitness value in this paper is the reciprocal of the classification accuracy, and the classification accuracy is the ratio of the same number of predicted values and actual values to the total number of samples. The classification accuracy expression is shown in Formula (11):

$$ACC = \frac{M}{N} \tag{11}$$

where M is the same number of predicted values and actual values, N is the total number of samples, and the fitness function is shown in Formula (12):

$$y = \frac{1}{ACC} \tag{12}$$

Combining the sparrow search algorithm and CNN-LSTM models, the SSA-CNN-LSTM model was constructed. The structure of the SSA-CNN-LSTM model is as follow:

(1) Data preprocessing: data labeling, data set division, data normalization, and data format conversion.
(2) SSA parameter initialization: setting the number of sparrows as n, the number of producers as PD, the number of sparrows sensing danger as SD, the safety threshold as ST, and the alarm value as R2.
(3) Calculating the fitness value, and updating the location of producer and scrounger.
(4) According to the anti-predation behavior, updating the location of the sparrow population.

(5) Inputting the data into the CNN network, the data through the CNN layer, batch normalization layer, activation function layer, and average pooling layer.
(6) The data enters the LSTM neural network and is inputted to the full connection layer and softmax layer through the LSTM layer.
(7) Output results.

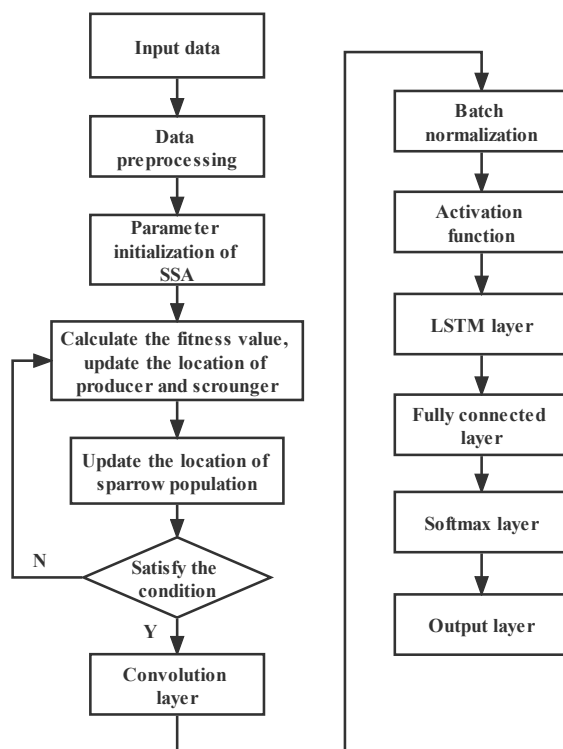Based on the above model structure construction steps, the SSA-CNN-LSTM model flow chart is shown in Figure 5:

Figure 5. Flow chart of the SSA-CNN-LSTM model.

The learning rate and the number of hidden layer nodes are important parameters of the neural network, which have a great influence on the training results. The learning rate controls the learning progress of the model. An excessive learning rate will make the network difficult to converge and linger near the optimal value. A learning rate which is too low will make the network converge very slowly and increase the time to find the optimal value. In this paper, the initial learning rate of the LSTM neural network was set to 0.0035. The number of hidden layer nodes has a certain influence on the performance of the model. Too many hidden layer nodes will increase the training time and increase the risk for the network to over-fit. Too few hidden layer nodes will make the network unable to learn successfully, increasing the number of training times and thus affecting the training accuracy. In this paper, the number of nodes in the two hidden layers of the LSTM model was set to 128 and 30, respectively. With the increase in the number of epochs, the network parameters were constantly updated to find the optimal value. The detailed parameter settings of the CNN and LSTM neural network are shown in Tables 2 and 3. SSA was used to optimize the learning rate of the model and the number of nodes in the second hidden layer of the LSTM model. The learning rate optimized by SSA was 0.0050725, and the number of nodes in the second hidden layer was 101. The decision variables in the model were the learning rate and the number of hidden layer nodes, and the constraints on the learning rate and the number of hidden layer nodes were set as upper bound (1×10^{-2}, 200) and lower bound (1×10^{-10}, 10). The parameters of SSA are shown in the Table 4. The data input process is shown in Figure 6:

Table 2. CNN parameters.

Parameters	Settings
CNN layers	1
Convolutional kernel size	1×1
Number of convolution kernels	32
Step	1
Activation function	ELU
Dropout	0.25

Table 3. LSTM neural network parameters.

Parameters	Settings
LSTM layers	2
Learning rate decline factor	0.5
Number of hidden layer nodes	30
Activation function	Adam
Epochs	800

Table 4. Sparrow search algorithm parameters.

Parameters	Settings
The number of populations	50
The maximum iterations	30
The number of producers	10
The number of scroungers	40
The sparrow that senses danger	40
Safety value	0.8
Lower limit range of parameters	$(1 \times 10^{-10}, 10)$
Upper limit range of parameters	$(1 \times 10^{-2}, 200)$
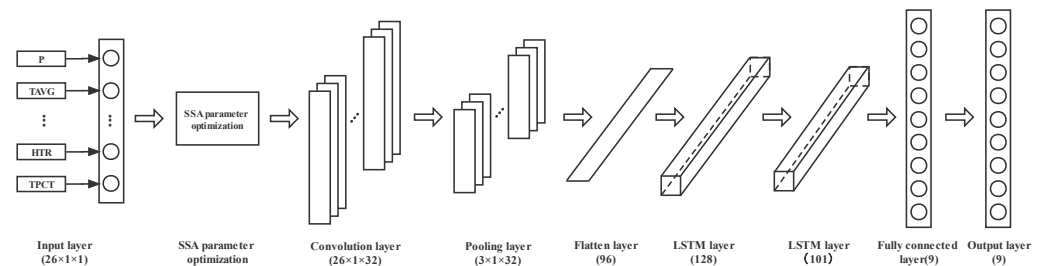
Figure 6. Architecture of the proposed regional SSA-CNN-LSTM model.

The data input process is shown in Figure 6. The input data was 26 parameters at each moment ($26 \times 1 \times 1$). After the convolution layer with 32 convolution kernels (1×1), the data became $26 \times 1 \times 32$. The data could learn the multi-dimensional features in the data through 32 convolution kernels (1×1). The data was then reduced to $3 \times 1 \times 32$ through the pooling layer with a pooling kernel of $10 \times 1 \times 1$. The multidimensional data was one-dimensionalized through the flatten layer, and then the fault category was outputted through the LSTM layer and the fully connected layer.

3. Experimental Data and Analysis

This paper focuses on the application of the SSA-CNN-LSTM model to the fault classification of nuclear power plants and compares the performance indicators of the SSA-CNN-LSTM model with the LSTM neural network and CNN-LSTM model. The experimental environment, and software and hardware configurations are shown in Table 5.

Table 5. Experimental conditions.

Name	Parameters
Simulation Software	Python 3.8.13
Operating system	Microsoft Win10
Hardware configuration	Intel(R) Core(TM) i7-1065G7 CPU @ 1.30 GHz 1.50 GHz

3.1. Experimental Data

In practical situations, data used in fault diagnosis are derived from sensor measurements. Due to the particularity of the nuclear power plant, operational data in accidents are scarce. Instead, data from a PCTRAN simulator was used as practical operation data to verify the fault diagnosis methods.

In this paper, the CPR1000 pressurized water reactor in the PCTRAN was selected as the object. One normal state and 8 accident conditions were simulated, as shown in Table 6. The fault was inserted at the 50th second of the PCTRAN operation and continued to run for 1000 s. The data included 92 parameters such as the containment pressure and hot leg temperature. From the 92 parameters, 26 parameters related to the fault were selected, as shown in Table 7. All data included 9 operating conditions, each with 26 parameters, running for 1000 s at a total of 245,700 data. Eighty percent of the data set was divided into the training dataset, and the remaining 20% into the testing dataset.

Table 6. Number of Accident Conditions.

Number	Accident Type
0	Normal operation
1	Loss of Coolant Accident (Hot Leg) (50 cm^2 break)
2	Loss of Coolant Accident (Hot Leg) (90 cm^2 break)
3	Steam Line Break Inside Containment (50 cm^2 break)
4	Steam Line Break Inside Containment (90 cm^2 break)
5	Loss of AC Power
6	Loss of Flow (Locked Rotor)
7	Steam Generator A Tube Rupture (50% of 1 full tube rupture)
8	Steam Generator A Tube Rupture (90% of 1 full tube rupture)

Table 7. Fault parameter names and symbols.

Serial Number	Parameter Name	Parameter Symbolic
1	Press RCS	P
2	Temperature RCS average	TAVG
3	Temperature Hot leg A	THA
4	Temperature Hot leg B	THB
5	Temperature Cold leg A	TCA
6	Temperature Cold leg B	TCB
7	Flow Reactor coolant loop A	WRCA
8	Flow Reactor coolant loop B	WRCB
9	Pressure Steam generator A	PSGA
10	Pressure Steam generator B	PSGB
11	Flow SG A feedwater	WFWA
12	Flow SG B feedwater	WFWB
13	Flow SG A steam	WSTA
14	Flow SG B steam	WSTB
15	Volume RCS liquid	VOL
16	Level Pressurizer	LVPZ
17	Flow Total ECCS	WECS
18	Power Total megawatt thermal	QMWT
19	Level SG A wide range	LSGA

Table 7. *Cont.*

Serial Number	Parameter Name	Parameter Symbolic
20	Level SG B wide range	LSGB
21	Level SG A narrow range	NSGA
22	Level SG B narrow range	NSGB
23	Temp Loop A subcooling margin	SCMA
24	Temp Loop B subcooling margin	SCMB
25	Power Pressurizer heater	HTR
26	Temp Peak clad	TPCT

3.2. Results Analysis

The accuracy and loss value of the LSTM neural network, CNN-LSTM model, and SSA-CNN-LSTM model were compared. The accuracy represents the proportion of the model that predicts the correct number out of the total. The loss value is the difference between the predicted value and the true value calculated by the loss function. The higher the accuracy, the lower the loss value, and therefore the better the model performance.

Figures 7 and 8 show the accuracy and loss value of the model training. The SSA-CNN-LSTM model achieved higher accuracy and a smaller loss value than the LSTM neural network and CNN-LSTM model in smaller epochs. The number of epochs refers to the number of training times for all training dataset. When the accuracy is close to 1, the loss value is close to 0 and the subsequent change is small, meaning the number of epochs can be considered as appropriate. Otherwise, it is necessary to continue to increase the number of epochs or adjust the network structure. When the number of epochs reached 800, the accuracy reached 0.9869, and the loss value was 0.0206. Therefore, the SSA-CNN-LSTM model had faster convergence speed, higher accuracy, and smaller loss value than those of the LSTM neural network and CNN-LSTM model. The training accuracy and test accuracy of the three models after 800 epochs are shown in Figure 9.
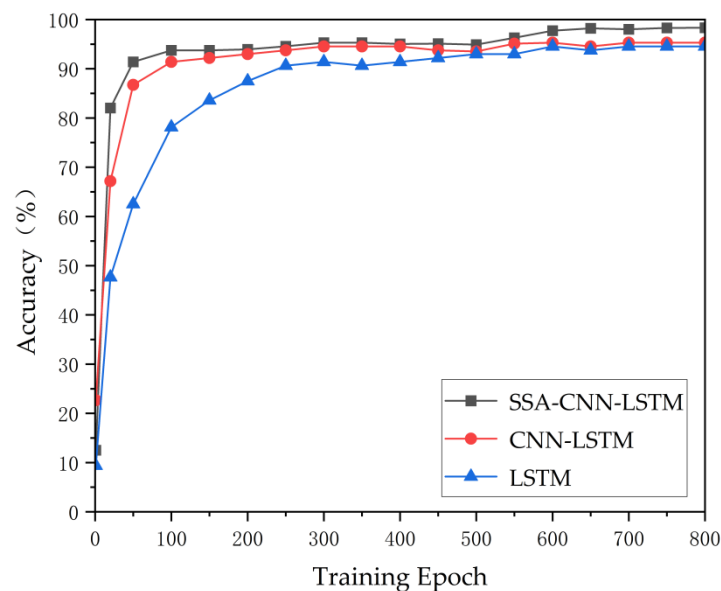

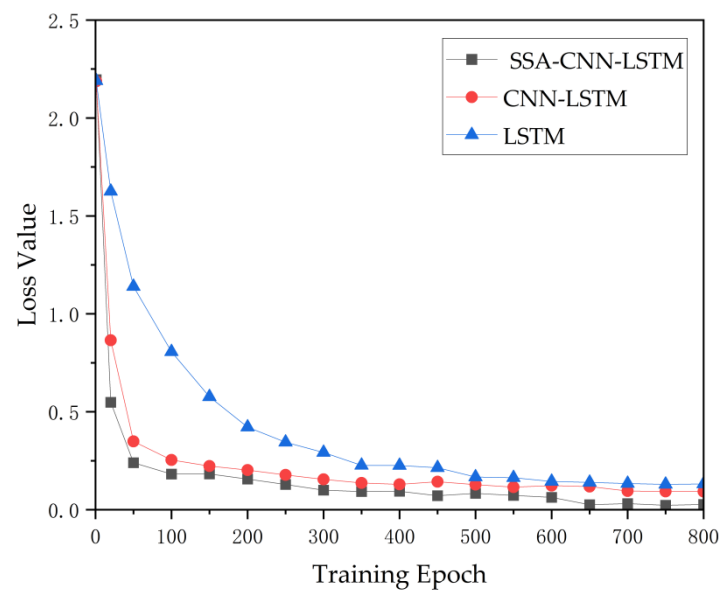**Figure 7.** Accuracy of the training dataset.

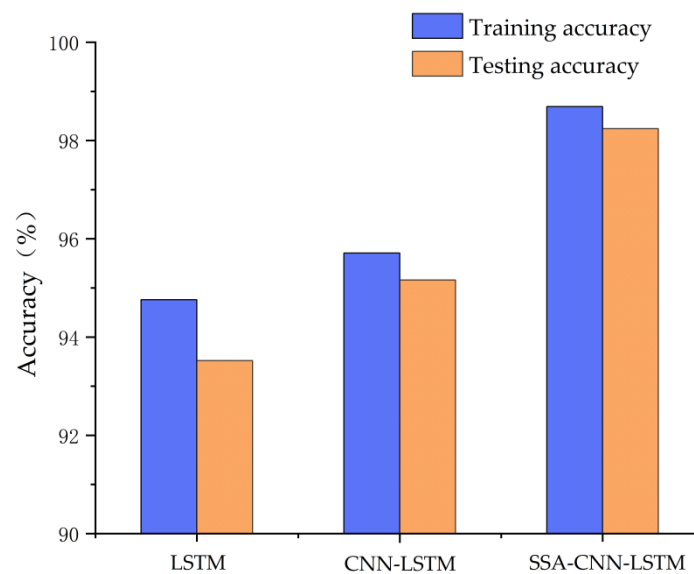**Figure 8.** Loss value of the training dataset.

Figure 9. Training and testing accuracy of the models.

3.3. Analysis of Confusion Matrix

The classification of each model in each fault category could be observed through the confusion matrix. The fewer the numbers on the non-diagonal line, the higher the classification accuracy. Figure 10 shows the confusion matrix of the testing dataset. It can be seen from the diagram that the SSA-CNN-LSTM model had a high accuracy of fault identification for 9 working conditions of the testing dataset, which were at 100.0%, 95.3%, 95.7%, 95.5%, 100.0%, 100.0%, 100%, 100%, 100%, and 100%.

By comparing the confusion matrix of the three models, it could be seen that the number of classification errors of fault 1, 2, and 8 were higher. After analysis, the original data of fault 1 and 2, and fault 7 and 8 were at different degrees of fault data of the same accident. Therefore, the accuracy of the model proposed in this paper was slightly worse than that of the different faults, but the overall effect was still good.

Select metrics based on the confusion matrix of the testing dataset classification were used. Precision and Recall are used to compare the diagnosis results of different methods. Precision P and recall R can be expressed as shown in Formulas (13) and (14):

$$P = \frac{TP}{TP + FP} \tag{13}$$

$$R = \frac{TP}{TP + FN} \tag{14}$$

where TP represents the number of correct samples predicted as correct samples; FP indicates the number of wrong samples predicted as correct samples; and FN indicates the number of correct samples predicted as wrong samples. The precision and recall of the LSTM neural network, CNN-LSTM model and SSA-CNN-LSTM model used in this paper are shown in Table 8:
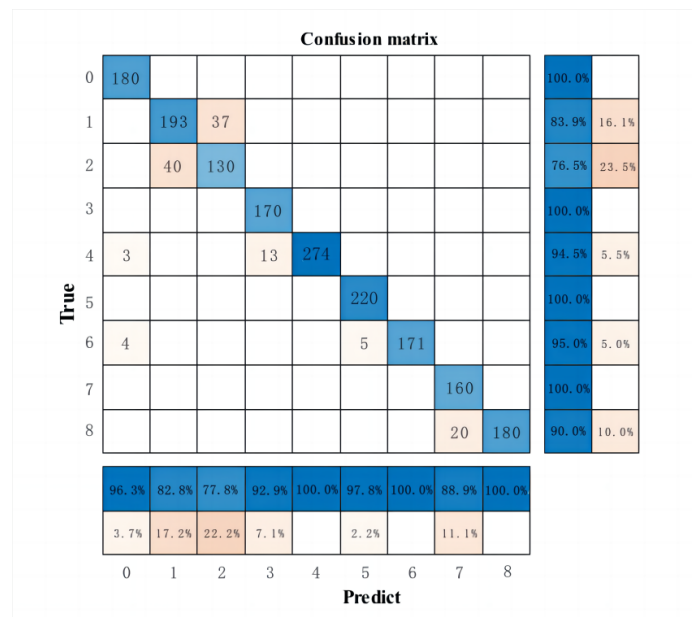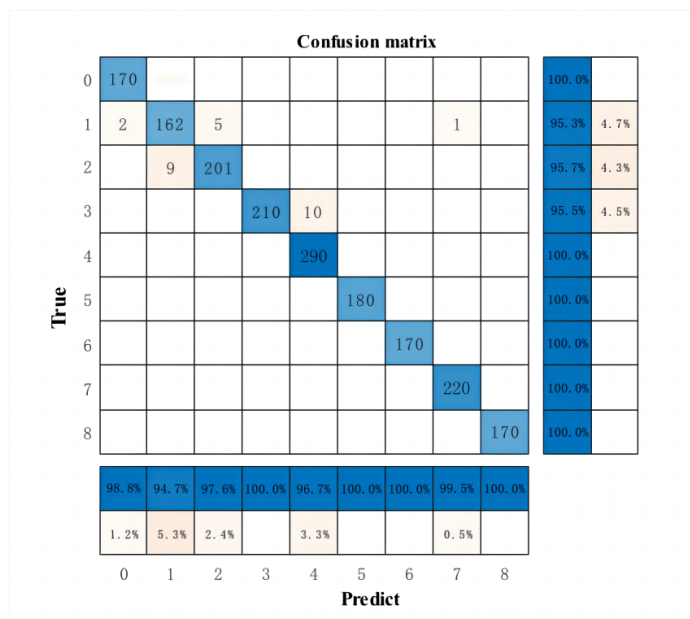
(a)
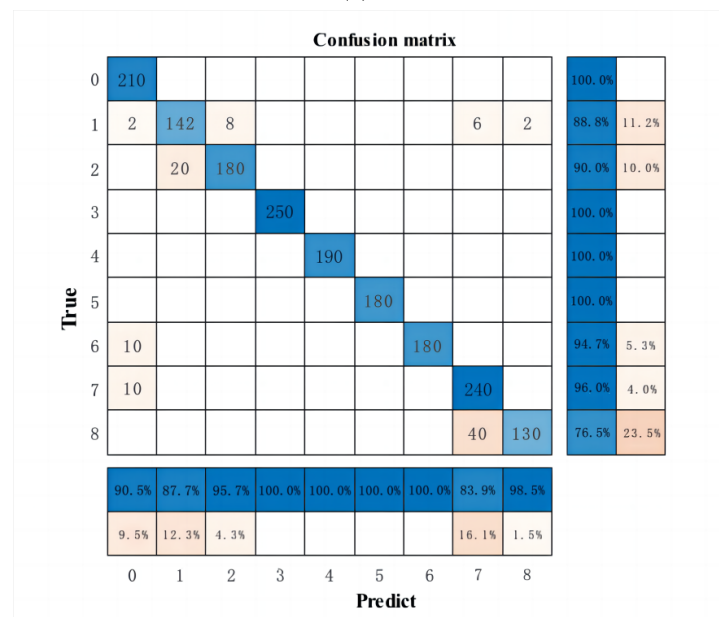
(b)

Figure 10. *Cont.*

(c)

Figure 10. Confusion matrix of the testing dataset. (**a**) Confusion matrix of testing dataset for the LSTM neural network, (**b**) Confusion matrix of testing dataset for the CNN-LSTM model, and (**c**) Confusion matrix of testing dataset for the SSA-CNN-LSTM model.

Table 8. Accuracy and recall rates of the models.

Fault Category	LSTM		CNN-LSTM		SSA-CNN-LSTM	
	Precision	Recall	Precision	Recall	Precision	Recall
0	100.0%	96.3%	100.0%	90.5%	100.0%	98.8%
1	83.9%	82.8%	88.8%	87.7%	95.3%	94.7%
2	76.5%	77.8%	90.0%	95.7%	95.7%	97.6%
3	100.0%	92.9%	100.0%	100.0%	95.5%	100.0%
4	94.5%	100.0%	100.0%	100.0%	100.0%	96.7%
5	100.0%	97.8%	100.0%	100.0%	100.0%	100.0%
6	95.0%	100.0%	94.7%	100.0%	100.0%	100.0%
7	100.0%	88.9%	96.0%	83.9%	100.0%	99.5%
8	90.0%	100.0%	76.5%	98.5%	100.0%	100.0%

It can be seen from the table that the precision and recall of the SSA-CNN-LSTM model for normal working conditions and 8 accident working conditions were generally higher than the LSTM neural network and CNN-LSTM model. The SSA-CNN-LSTM model only had a lower precision for fault 3 and a lower recall for fault 4 in contrast to the other two models, and the remaining metrics were better than the other two models. It could be seen that the introduction of SSA can improve the overall precision and recall of the model.

4. Conclusions

This paper proposes a fault diagnosis method for nuclear power plants based on the SSA-CNN-LSTM model. Firstly, the CNN was used to extract features, and then the data after feature extraction was sent to the LSTM neural network to mine the time series features of data. Finally, the SSA was used to optimize the parameters of the LSTM neural network to obtain the SSA-CNN-LSTM model. In this study, the SSA and CNN-LSTM model were innovatively used in nuclear power plant fault diagnosis problems with good results.

The SSA-CNN-LSTM model was validated using the run data of a PCTRAN and compared with the LSTM and CNN-LSTM models. The experimental results showed that the fault identification accuracy of the LSTM and CNN-LSTM models were 95.16%

and 93.52%, respectively, and the fault identification accuracy of the SSA-CNN-LSTM model with the addition of the optimization algorithm was improved to 98.24%. All three models in the paper are capable of classifying nuclear power plant faults, but there will be differences in the accuracy of the classification. The SSA-CNN-LSTM model had the highest accuracy, which proves that the CNN-LSTM has a higher classification accuracy compared with a single model and that the SSA has a better effect on the optimization of the model.

Compared with the traditional machine learning model, the SSA-CNN-LSTM model proposed in this paper has the ability to process the complex data of nuclear power plants, can dig deeper into the timing characteristics, and has higher prediction accuracy in the fault diagnosis of nuclear power plants. When an accident occurs in a nuclear power plant, the SSA-CNN-LSTM model can determine the type of accident 0.22 s after the accident, which is of great significance for helping operators quickly identify faults, take corresponding measures in time, and improve the safety of nuclear power plant operation. However, this method has limitations. The accuracy of classification results will be reduced for unknown or untrained incidents. In the case of a small size in sample data, the model may not be able to fully learn the data features, and the accuracy may also be reduced. In addition, the training time of the SSA-CNN-LSTM model is long and is significantly longer than that of the LSTM neural network and CNN-LSTM model. Future research can improve models for these problems and further develop fault diagnosis models for nuclear power plants.

References

1. Sadekin, S.; Zaman, S.; Mahfuz, M.; Sarker, R. Nuclear power as foundation of a clean energy future: A review. *Energy Procedia* **2019**, *160*, 513–518. [CrossRef]
2. Mu, R.; Zuo, J.; Yuan, X. China's approach to nuclear safety—From the perspective of policy and institutional system. *Energy Policy* **2015**, *76*, 161–172. [CrossRef]
3. Mo, K.; Lee, S.J.; Seong, P.H. A dynamic neural network aggregation model for transient diagnosis in nuclear power plants. *Prog. Nucl. Energy* **2007**, *49*, 262–272. [CrossRef]
4. Lee, U.; Lee, C.; Kim, M.; Kim, H.R. Analysis of the influence of nuclear facilities on environmental radiation by monitoring the highest nuclear power plant density region. *Nucl. Eng. Technol.* **2019**, *51*, 1626–1632. [CrossRef]
5. Zubair, M.; Ahmed, E.; Hartanto, D. Comparison of different glass materials to protect the operators from gamma-rays in the PET using MCNP code. *Radiat. Phys. Chem.* **2022**, *190*, 109818. [CrossRef]
6. Ur, R.K.; Heo, G.; Kim, M.C.; Zubair, M. Formulation and Reliability Feature Analysis of Analog, Digital and Hybrid I&C Architectures for Research Reactors. In Proceedings of the 22nd International Conference on Nuclear Engineering, Prague, Czech Republic, 7–11 July 2014.
7. Cao, H.; Sun, P.; Zhao, L. PCA-SVM method with sliding window for online fault diagnosis of a small pressurized water reactor. *Ann. Nucl. Energy* **2022**, *171*, 109036. [CrossRef]
8. Zio, E. A support vector machine integrated system for the classification of operation anomalies in nuclear components and systems. *Reliab. Eng. Syst. Saf.* **2007**, *92*, 593–600.
9. Mu, Y.; Xia, H.; Liu., Y. Fault diagnosis method for nuclear power plant based on decision tree and neighborhood rough sets. *At. Energy Sci. Technol.* **2011**, *45*, 44.
10. Xu, J.; Chen, W.; Tang, Y. Study on fault diagnosis in nuclear power plant based on Rough Sets and Support Vector Machine. *Nucl. Power Eng.* **2009**, *30*, 51–54.

11. Zhang, Q.; An, X.; Gu, J.; Zhao, B.; Xu, D.; Xi, S. Application of FBOLES—A prototype expert system for fault diagnosis in nuclear power plants. *Reliab. Eng. Syst. Saf.* **1994**, *44*, 225–235. [CrossRef]
12. Mwangi, M.A.; Yong-Kuo, L.; Ochieng, A.S. Small Break Loss of Coolant Accident (SB-LOCA) fault diagnosis using Adaptive Neuro-Fuzzy Inference System (ANFIS). *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *675*, 012034. [CrossRef]
13. Akyol, S.; Alatas, B. Plant intelligence based metaheuristic optimization algorithms. *Artif. Intell. Rev.* **2017**, *47*, 417–462. [CrossRef]
14. Mehrabian, A.R.; Lucas, C. A novel numerical optimization algorithm inspired from weed colonization. *Ecol. Inform.* **2006**, *1*, 355–366. [CrossRef]
15. Murase, H. Finite element inverse analysis using a photosynthetic algorithm. *Comput. Electron. Agric.* **2000**, *29*, 115–123. [CrossRef]
16. Yu, S.; Zhu, S.; Ma, Y.; Mao, D. A variable step size firefly algorithm for numerical optimization. *Appl. Math. Comput.* **2015**, *263*, 214–220. [CrossRef]
17. Tilahun, S.L.; Ngnotchouye, J.M.T. Firefly algorithm for discrete optimization problems: A survey. *KSCE J. Civ. Eng.* **2017**, *21*, 535–545. [CrossRef]
18. Yang, X.S.; He, X. Bat algorithm: Literature review and applications. *Int. J. Bio-Inspired Comput.* **2013**, *5*, 141–149. [CrossRef]
19. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [CrossRef]
20. Mirjalili, S. The ant lion optimizer. *Adv. Eng. Softw.* **2015**, *83*, 80–98. [CrossRef]
21. Mirjalili, S.; Lewi, A. The whale optimization algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [CrossRef]
22. Kirkpatrick, S.; Gelatt, C.D., Jr.; Vecchi, M.P. Optimization by simulated annealing. *Science* **1983**, *220*, 671–680. [CrossRef]
23. Hatamlou, A. Black hole: A new heuristic optimization approach for data clustering. *Inf. Sci.* **2013**, *222*, 175–184. [CrossRef]
24. Atashpaz-Gargari, E.; Lucas, C. Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition. In Proceedings of the 2007 IEEE Congress on Evolutionary Computation, Singapore, 25–28 September 2007; pp. 4661–4667.
25. Geem, Z.W.; Kim, J.H.; Loganathan, G.V. A new heuristic optimization algorithm: Harmony search. *Simulation* **2001**, *76*, 60–68. [CrossRef]
26. Lam, A.Y.S.; Li, V.O.K. Chemical-reaction-inspired metaheuristic for optimization. *IEEE Trans. Evol. Comput.* **2009**, *14*, 381–399. [CrossRef]
27. de Lima, A.M.M.; Schirru, R.; da Silva, F.C.; Medeiros, J.A.C.C. A nuclear reactor core fuel reload optimization using artificial ant colony connective networks. *Ann. Nucl. Energy* **2008**, *35*, 1606–1612. [CrossRef]
28. Khoshahval, F.; Minuchehr, H.; Zolfaghari, A. Performance evaluation of PSO and GA in PWR core loading pattern optimization. *Nucl. Eng. Des.* **2011**, *241*, 799–808. [CrossRef]
29. Zhou, D.X. Theory of deep convolutional neural networks: Downsampling. *Neural Netw.* **2020**, *124*, 319–327. [CrossRef] [PubMed]
30. Xue, J.; Shen, B. A novel swarm intelligence optimization approach: Sparrow search algorithm. *Syst. Sci. Control Eng.* **2020**, *8*, 22–34. [CrossRef]