# An Enhanced FSO-BPNN Framework for Anomaly Detection and Early Warning in Power System Monitoring

Na Li*, Guanghua Yang, Yuexiao Liu, Xiangyu Lu, Zhu Tang
State Grid Beijing Electric Power Company, Beijing,100032, China
E-mail：diance003@126.com
*Corresponding Author

**Keywords:** anomaly detection (AD), internet of things (IOT), monitoring, neural network, power system (PS) smart grid, predictive maintenance

**Recieved:** May 7, 2025

*The increasing complexity of contemporary power networks necessitates the development of enhanced early warning systems and intelligent monitoring to ensure stability and operational efficiency. Traditional approaches to risk prevention and predictive maintenance often fail due to limitations in identifying real-time abnormalities and adapting to dynamic system characteristics. To address these issues, the present research proposes an improved fish swarm optimization with Backpropagation Neural Network (IFSO-BPNN) for anomaly detection (AD) and fault detection (FD) early warning in power system (PS) monitoring that integrates an IFSO algorithm with a BPNN. The major goal is to increase the accuracy of AD and FD in smart grids by utilizing deep learning (DL) and optimization approaches. The IFSO method integrates adaptive weighting and behavioral dynamics into classic fish swarm optimization, improving overall search capabilities. By tweaking BPNN parameters using IFSO, the model achieves higher convergence rates and improved classification accuracy. The assessment dataset was compiled usingInternet of Things (IoT) sensors and pan/tilt camera-based surveillance systems at Beijing power plants, with preprocessing techniques such as min-max normalization and feature extraction using Independent Component Analysis (ICA) to improve model performance. Resultsfrom experiments show that the IFSO-BPNN model outperforms standard algorithms with an accuracy ofFD99.98% and AD 0.9980. These findings illustrate the system's capacity to detect anomalies quickly and perform preventive maintenance. The proposed method, which combines swarm intelligence with neural networks, helps to construct smarter, more robust power grids capable of meeting future energy demands with lower failure risks.*

*Povzetek: Za odkrivanje napak (FD) in nepravilnosti (AD) v nadzoru elektroenergetskega sistema je razvit IFSO-BPNN (Izboljšana optimizacija jata rib in BPNN). Model izboljša kvaliteto z optimizacijo parametrov BPNN z IFSO, kar omogoča hitro zgodnje opozarjanje in prediktivno vzdrževanje.*

## 1 Introduction

Artificial intelligence (AI), big data, and deep learning (DL) revolutionize power systems (PS) by enhancing feature modeling, control, and fault diagnosis; these are presenting recent advances and applications in monitoring and performance analysis [1]. The expansion ofPS is hindered by growing power demand and environmental objectives, which present challenges for transmission capacity and distance. Advanced, sustainable energy solutions are being used to achieve carbon peaking and neutrality [2].Reconstruction errors and thresholding are used in AD(AD) to minimize false alarms and isolate fault areas by training a model to learn typical system behaviorin an unsupervised manner [3].Approximately 70% of energy is produced by thermal power plants; new large-capacity units (600–1000+ MW) improve operating efficiency but make system coupling and integration more difficult [4].Real-time data collection and analysis of electrical characteristics is part of PS monitoring, used to ensure system stability, identify problems, improve performance, and assist in decision-making for dependable and effective power grid operation [5].As demonstrated by the arctic sky tragedy, the expansion of the cruise industry needs advanced, dependable PS to avoid blackouts, which endanger public safety, the environment, financial stability, and reputation [6].Potential false alarms, reliance on data quality, difficulty identifying new abnormalities, computational complexity, difficulties with real-time implementation, and threshold setting are some drawbacks of AD and early warning in PS [7].

### 1.1 Aim and contribution of the research

The aim of the research is to develop a new method, improved fish swarm optimization with Backpropagation Neural Network (IFSO-BPNN), for detecting anomalies and faultsin PS by integrating BPNN and IFSO algorithms. The goal is to increase the accuracy and efficiency of AD and fault detection (FD) in smart grids while also enabling proactive maintenance. The research's key contributions include the following:

- ***IFSO Algorithm***: Improves the global search capability and adaptive weighting of classic Fish Swarm Optimization, resulting in less convergence time and higher classification accuracy in anomaly and fault identification.
- ***BPNN Optimization***: IFSO is used to optimize BPNN parameters, which results in quicker convergence and greater classification accuracy for real-time AD and FD.
- ***Advanced-Data Preprocessing:*** Uses min-max normalization and Independent Component Analysis (ICA) for feature extraction, improving the model's performance in power system monitoring by efficiently preprocessing Internet of Things (IoT) sensor and surveillance system data.

The next phase (phase 2) clearly explains the existing research about ADand early warning in PS monitoring. Phase 3 presents the methodology, Phase 4 provides the result and discussion of existing vs proposed method, and Phase 5 deliversthe conclusion.

## 2 Related works

The aim of the research [8] was to increase the dependability of seismic stations. For reliable power failure prediction, the SeismoGuard Ensemble, which comprises random forest (RF), support vector machine (SVM), k-nearest neighbors (KNN), and logistic regression (LR), along with IoT monitoring, was used. Results demonstrate that the approach attained 90% accuracy and increased dependability. The dataset's reach was restricted; however, the data contains long-term testing with wider generalization across various situations. A combination of elliptic curve cryptography (ECC)-based token control with deep reinforcement learning (DRL)-based sleep scheduling was used for secure and adaptive power management under possible threat conditions in order to improve the security and energy efficiency of wireless sensor networks (WSNs) [9]. The approach achieved a 15% increase in energy efficiency and a 20.01% power reduction. While simulation-based outcomes were validated, more verification was required for scalability and real-world implementation under various attack types.

Following data cleaning and feature extraction, supervisory control and data acquisition (SCADA)were processed using aConvolutional neural network - bidirectional gated recurrent unit (CNN-BiGRU) with attention to identify wind turbine faults [10]. Accurate FD in actual wind farms was accomplished; however, it was constrained by the generalizability of the data source and the possibility of overfitting to particular turbine models. The monitoring of wind turbine health was enhanced by utilizing mutual information to determine essential parameters, support vector regression (SVR) for thresholding, and long short-term memory -autoencoder (LSTM-AE) for AD [11]. The outcome demonstrated precise AD and successful identification of crucial parameters. Real-time monitoring settings could show a decline in performance due to noisy data or inadequate temporal information. To optimize the monitoring and security of smart hospitals, machine learning (ML) and edge-based advertising on Contiki Coojawere applied to identify IoT network intrusions and e-health incidents [12]. The system was successful in identifying cyberattacks and e-health events, but it was very dependent on the reality of the simulated data, which could not work effectively with complex or novel attack patterns.

Abnormalities in wind turbines were discovered and accurately analyzedutilizing a combination of methods. Local outlier factor (LOF) and adaptive K-means for preprocessing, Extreme Gradient Boosting (XGBoost) for diagnosis, and long short-term memory-stacked denoising autoencoder (LSTM-SDAE) for feature extractionwere employed [13]. The technique increased wind turbine dependability by efficiently identifying and diagnosing problems in real-time utilizing SCADA data. Performance was dependent on the caliber of preprocessing and could be hampered by noisy data or hidden anomalies. The research created an early warning system that incorporates meteorological data to enhance PS dependability and proactively reduce atmospheric dangers [14]. The technology enhanced defect detection and prevented outages during severe weather; however, its performance depended on data quality and erratic weather patterns. The advancements in battery electric vehicle (BEV) technology, platforms, charging, and monitoring were examined to address issues regarding safety, charging, and range in new energy cars [15]. Although cutting-edge platforms and safety features dominate the BEV industry, however, there were issues with battery lifecycle safety, charging simplicity, and weather adaptation. The PS load margin was determined by utilizing an artificial neural network (ANN) trained on phasor measurement unit (PMU) data and model simulations to ensure voltage and small-signal stability [16]. An ANN's ability to anticipate load margin effectively cannot exceed a dependence on the quality of PMU data and model assumptions in actual systems. To increase safety in nuclear-powered marine operations, developments in ship nuclear power machinery (SNPM) design, fault diagnostics, and risk assessmentwere evaluated [17]. Design enhancements and investigation spaces were identified, and an integrated risk framework was suggested; however, knowledge remains limited and needs to be verified. Table 1 provides the related works summary table.

Table 1: Comparative Summary of the related works

| Reference | Methods | Results | Limitations |
|---|---|---|---|
| Duet al. [8] | SeismoGuard Ensemble (RF, SVM, KNN, LR) + IoT monitoring | Achieved 90% accuracy, improved dependability of seismic stations | Limited dataset coverage; needs generalization and broader testing |
| Qinet al.[9] | ECC token control + DRL-based sleep scheduling for WSN | 15% energy efficiency gain, 20.01% power reduction | Simulation-based only; real-world scalability and threat resilience not verified |
| Xianget al.[10] | SCADA data + CNN-BiGRU + attention mechanism | Accurate wind turbine FD in real wind farms | Data source generalizability is limited; overfitting risk to specific turbine models |
| Chen et al. [11] | Mutual information + SVR for thresholds + LSTM-AE for anomaly detection | Accurate anomaly detection; key parameters identified | Real-time performance could degrade under noisy or incomplete data |
| Said et al. [12] | ML + edge-based intrusion detection on Contiki Cooja for smart hospitals | Identified e-health events and IoT network intrusions accurately | Simulated data could fail under real, complex attack patterns |
| Zhang et al. [13] | LOF + adaptive K-means preprocessing + XGBoost + LSTM-SDAE | Real-time, accurate ADand diagnosis in wind turbines | Sensitive to data quality; hidden anomaly types may be missed |
| Božičeket al.[14] | Early warning system using meteorological data | Prevented outages and improved detection during extreme weather | Dependent on weather unpredictability and data quality |
| He et al. [15] | BEV platform, charging/swapping stations, and monitoring platform | Technological dominance and safety improvements in the BEV market | Issues remain in battery safety, weather adaptability, and charging ease |
| Bento et al. [16] | ANN trained on PMU data + model-based simulation | Accurate load margin prediction ensuring voltage and small-signal stability | Performance hinges on PMU data and assumptions in simulation models |
| Adumene et al. [17] | SNPM designs + fault diagnosis + risk assessment | hybrid risk framework; identified design progress | Incomplete knowledge base; needs validation and framework integration |

## 2.1 Research gap

The method additionally solves past techniques' drawbacks, such as restricted data generalization, overfitting, simulation reliance, and data quality sensitivity. The proposed approach, IFSO-BPNN, provides a scalable, real-time solution for proactive maintenance and problem detection in complex, large-scale power networks. The research fills a gap by merging an IFSO method with a BPNN for PS anomaly and fault identification. Compared to earlier techniques, this approach improves accuracy, convergence speed, and FD resilience, especially in noisy situations.

# 3 Research methodology

This section discusses IoT sensor-based data collection in PS and introduces the IFSO-BPNN approach for anomaly and fault identification, as well as early warning in PS monitoring. Figure 1 shows the methodology flow, which includes data pretreatment, feature extraction, and model optimization.

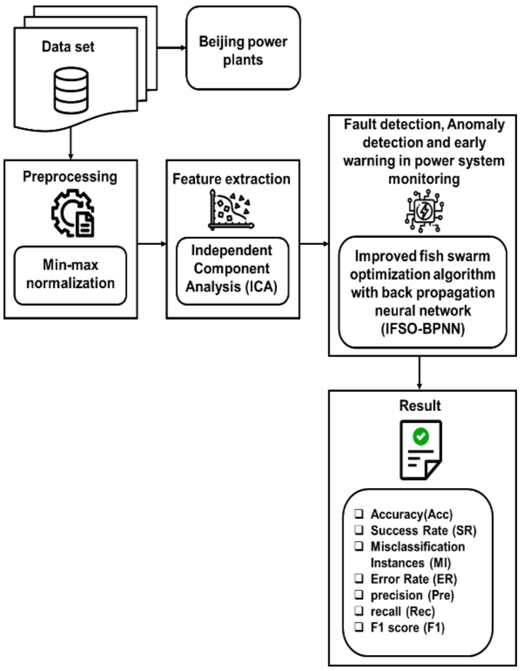


Figure 1: Flow of the proposed method

## 3.1 Data collection

The system configuration includes a pan/tilt integrated camera, a series of local storage DVR hosts, a 1-terabyte dedicated hard disk, and equipment from major domestic video equipment manufacturers. A wireless networking module is an important element that allows direct connection across 4G or 5G wireless networks. The research is centered on power stations surrounding Beijing, where the distribution stations lack wired networks and must communicate over wireless networks. To achieve that, on-site terminal equipment is required to access different network types at the distribution station, such as 2G/3G/4G, GSM, CDMA, and wired networks. Many of these stations are found in basements. In the event of a severed wireless connection between the station and the platform, short messages transmitted to the terminal equipment at the distribution station allow for simple permission and re-establishment of communication. The data were split into an 8:2 ratio, 80% for training, and 20% for testing dataset.

## 3.2 Data preprocessing via min-max normalization

Min-max normalization is a common method used for numerical sensor and camera data from Beijing power plants to scale characteristics between 0 and 1, in which the values of a feature are translated into a preset range, usually [0-1]. The method retains data connections, hence being suitable for a wide range of ML applications. The transformation is carried out using the following Equation (1).

$$X_{new} = \frac{x=min}{max(x)-min(x)} \quad (1)$$

$X_{new}$ = The adjusted value obtained after scaling the data $X$ = outdated value, $max(x)$= dataset's highest possible value.$min(x)$ = dataset's lowest possible value. The normalizing technique improves AD and FD in PS monitoring by ensuring that all data points have a consistent scale, which increases predictive model accuracy.

## 3.3 Feature extraction using independent component analysis (ICA)

ICA is a current statistical technique that attempts to break down observable data into statistically independent components. The ICA was used on sensor and surveillance data to reduce dimensionality and extract essential features, which improved the IFSO-BPNN model's capacity to detect abnormalities in PS monitoring as a linear mixture of independent components, expressed as follows in Equation (2).

$$y = B.T \quad (2)$$

Where: $y$ represents the observed data vector, $B$denotes the mixing matrix, and $T$ denotes the separate components. In ICA, components are assumed to be statistically independent and non-Gaussian, with a square and unknown mixing matrix$B$. To extract the components, calculate the inverse $X$ of matrix $B$ as follows in Equation (3).

$$T = X.y \quad (3)$$

ICA divides data into statistically independent components, helping in AD and FD in PS. While the technique does not give direct variance or ordered data, the enhanced sparsity-based technique improves feature extraction and speeds up convergence for real-time applications such as early warning systems. ICA has been widely applied in disciplines like face recognition and dimensionality reduction. PS monitoring, which extracts essential characteristics from sensor data, catches complicated, non-Gaussian patterns that standard approaches typically overlook, resulting in improved AD, FD and maintenance efficiency.

## 3.4 Detection and early warning in PS monitoring using improved fish swarm optimization with backpropagation neural network (IFSO-BPNN)

The IFSO-BPNN enhances AD and FD in PSby optimizing BPNN parameters with the IFSO algorithm, increasing classification accuracy, and allowing for real-time predictive maintenance. Figure 2 displays the proposed method's flow diagram for power system monitoring.

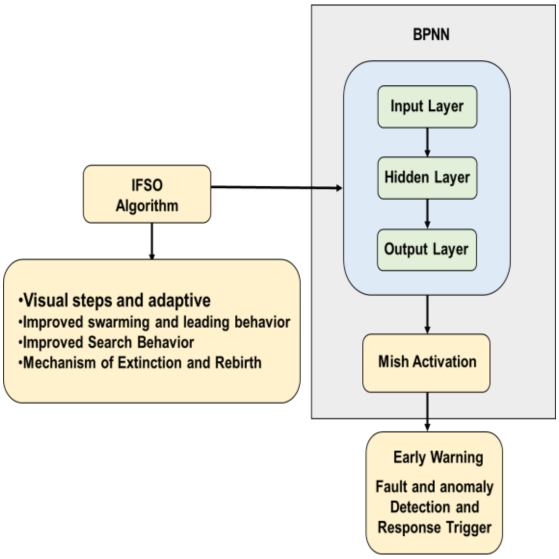


Figure 2: Flow diagram for the proposed method.

### 3.4.1 Back-propagation neural network (BPNN)

The BPNN is a multi-layer feed-forward artificial neural network designed to identify anomalies in PS. The architecture consists of an input layer, one or more hidden layers, and an output layer. Sensor readings, system performance measurements, and ambient parameters are all sent into the input layer. The hidden layers discover complicated patterns in the data, whereas the output layer anticipates anomalies and faults such as system malfunctions or failures. Each neuron's output is defined by applying an activation function to the weighted sum of inputs in Equation (4).

$$x = \sum_{j=1}^{n} z_j . y_i + a \quad (4)$$

Where $y_i$is the input, $z_j$is the weight, $a$is the bias, and $\sigma$ ($\cdot$) is the exponential activation function (TanhExp) $f(x) = x.\tanh(e^x)$, generally mish activation function $tanh$. The Mish function is smooth and comparable to TanhExp. The formula is provided as follows. $F(x)x.tanh(softplus(x))$. where softplus function$f(x) = \log(1 + e^x)$.Mish is a self-regulatory activation that improves accuracy and generalization instead of standard function. The process is smooth and non-monotonic, allowing for modest negative outputs while retaining strong positive flow, avoiding problems like dead neurons in ReLU.$x$: Input to the neuron. $Soft\,plus$ (x): A smooth variant of ReLU.$tanh(\cdot)$: Implements smooth limiting behavior for high input values. Data from the power system is collected, standardized, and sent to the network for training. Normalization guarantees that each input feature contributes evenly to model training. During forward propagation, input data is transferred through the layers as the model produces predictions. Backpropagation then changes the weights and biases depending on the loss function, which is commonly Mean Squared Error (MSE) and computed as follows in Equation (5).

$$MSE = \frac{1}{N}\sum_{j=1}^{N}\left(x_{pred} - x_{actual}\right)^2 \quad (5)$$

To improve the model's capacity to detect anomalies and fault, increase system dependability, and provide early alerts for proactive PS repair.

***Loss function:***

In the PS anomaly and fault detection, the loss function is critical for reducing prediction errors and improving model parameters. The BPNN's output layer computes the error between the expected output and the actual observed detection using the MSE and an appropriate activation function. The error gradient of each neuron in the output layer could be computed as follows in Equation (6).

$$\delta_{out} = \left(x_{pred} - x_{true}\right).\sigma'(w) \quad (6)$$

$x_{pred}$ : predicted output (anomaly, and fault score). $x_{true}$: True label (0 for no abnormality and 1 for anomaly). $\sigma'(w)$ is the derivative of the activation function for the neuron's input $w$.The gradient of the hidden layers is affected primarily by the output error, but also by the gradients of the following layers. The gradient of a hidden layer neuron $G_j$ could be calculated using the chain rule in Equation (7).

$$\delta_{hidden} = \sum_i z_{j.i}.\delta_i.\sigma'(w_j) \quad (7)$$

$\delta_{hidden}$: Error gradient for a hidden layer neuron. $z_{j.i}$: Weight coupling hidden layer cell $G_j$ with output neurons. $\delta_i$: The error gradient of the output neuron.$\sigma'(w_j)$: Derivative of the activation function for the buried layer input $w_j$.Gradient descent is used to update weights and biases during training to minimize the loss function. The rules for updating the weights $(z)$ and biases $(a)$ in each round are as follows in Equations (8-9).

$$z^{(n+1)} = z^{(n)} - \eta.\frac{\partial P}{\partial z} \quad (8)$$
$$a^{(n+1)} = a^{(n)} - \eta.\frac{\partial P}{\partial a} \quad (9)$$

The current weights and biases at iteration $n$ are denoted by $z^{(n)}$and$a^{(n)}$. The learning rate ($\eta$) is a hyperparameter that controls the step size. The gradients of the loss function about weights and biases are $\partial\frac{\partial P}{\partial z}$and$\frac{\partial P}{\partial a}$, respectively. The learning rate $\eta$ adjusts the model's weights and biases to reduce prediction errors for ADinPS.

### 3.4.2 Improved fish swarm optimization (IFSO)

FSO was selected over PSO, GA, and DE because of its greater global search capabilities and adaptive behavior, which improve convergence and classification accuracy in AD and FD. An IFSO is proposed to increase detection accuracy and convergence speed. For balanced exploration and exploitation, the system incorporates adaptive control overstep size and visual field, which shrinks with iterations. By eliminating default search behaviors and crowding conditions, swarming and following techniques are improved. Fish retry with modified settings when an improved solution is discovered. To preserve the quality of global optimization, an extinction-regeneration system removes the most susceptible fish and replaces it with a more suitable one. This improved method efficiently optimizes BPNN parameters for AD and FD in PS.

The classic Fish Swarm Algorithm (FSA) has fixed visual and step sizes, which can hinder convergence. To improve AD performance, an adaptive piecewise function is proposed to gradually decrease visual and step sizes with iterations, finding a balance between speed and accuracy. Step$SS$ $(iter)$ and adaptive $V(iter)$ are defined as follows in Equations (10-11).

$$V(iter) = \text{int}\left(max_v \times \left(\frac{\log(min_v/max_v)}{\log(max_{gen})}\right)^{iter}\right) (10)$$
$$SS(iter) = \text{int}\left(max_s \times \left(\frac{\log(min_s/max_s)}{\log(max_{gen})}\right)^{iter}\right) (11)$$

$V(iter)$: The artificial fish's field of vision at iteration iter. $SS(iter)$: The maximum step the fish can take during iteration. $max_v$: Step size and initial (maximum) visual range. $min_v$: The smallest step size and visual range for efficient searching. The maximum number of iterations is$max_{gen}$. $iter$: The number of the current iteration. For discrete issues, $int(\dots)$ rounds values to integers. Values are rounded to integers, with a minimum step and visual sizes set to 1 for discrete issues such as attribute reduction in Equations (10-11); both the visual and step sizes use an exponential decrease from maximum to minimum across iterations, allowing for quick global search at the beginning and accurate local search at the final stage. The provided AD, and FD framework's convergence and detection accuracy are enhanced by the adaptive technique.

The artificial fish swarm algorithm (AFSA) uses swarming and following behaviors to determine convergence speed. However, narrow distances can cause local optima and delayed convergence. Randomization changes swimming's step size to prevent premature convergence. The algorithm focuses on determining the optimal position of fake fish for efficient attribute reduction, and eliminates search behavior to save execution time. The enhanced swarming and subsequent behaviors are defined as Equations (12-13).

$$Y_{next} = Y_j + step \times (Y_d - Y_j) if\ G(Y_d) > G(Y_j) \quad (12)$$
$$Y_j = Y_d\ if\ G(Y_d) > G(Y_j) \quad (13)$$

$Y_{next}$: The fake fish's next position.$Y_j$: The fake fish's current location.$Y_d$: The position of the swarm's center.$step$: The step size for movement is determined by a random component. $G(Y_d)$: The fitness value at the center position. $G(Y_j)$: The fish's fitness value at that present location. These changes improve the algorithm's efficiency, resulting in faster convergence and higher performance.

Improved Search Behavior**:** In the AFSA, searching for behavior entails exploring the available domain to discover alternatives. The number of tries has a significant impact on search efficiency, frequently resulting in premature or inefficient searches. To solve these things, extend the viewing field when no superior location is discovered after a certain number of difficulties. When a suitable place is located, the fish takes one step towards that, with a maximum step size of $step_{new} = 2 \times step$. Without false, the fish moves randomly. IFSO's capacity was improved to efficiently tune BPNN parameters, hence increasing accuracy and convergence in PS anomaly and fault detection.

Mechanism of Extinction and Rebirth: The algorithm uses an extinction mechanism to remove the least suitable fish, enhancing swarm adaptability but decreasing swarm size and randomness. A regeneration mechanism is then included to restore swarm size by regenerating highly adaptable fish, ensuring resilience and enhancing efficiency by shortening iteration durations while maintaining high fitness levels. The IFSO-BPNN approach attempts to discover and detect deviations in PSmore efficiently by optimizing neural network parameters, assuring faster convergence, and improving prediction accuracy for proactive maintenance. Algorithm 1 displays IFSO-BPNN.

Algorithm 1: IFSO-BPNN

Step 1: Initialize the BPNN parameters
Initialize BPNN with input layer, hidden layers, and output layer
Set learning rate η and number of iterations max_iter
Step 2: Initialize the Fish Swarm Optimization (IFSO) parameters
Initialize fish swarm population size, maximum visual field (max_v), and step size (max_s)
Set the minimum values for visual field (min_v) and step size (min_s)
Step 3: Data Preprocessing
Preprocess data:
  Normalize sensor readings using min-max normalization
  Perform feature extraction using Independent Component Analysis (ICA)
Step 4: Training the BPNN with IFSO optimization

```
for each iteration in range(max_iter):
    for each fish in the swarm:
        visual = V(iter)
        step = SS(iter)
        if G(Y_d) > G(Y_j):
Y_j = Y_d
    for each fish in the swarm:
BPNN.weights = optimize_with_fish_swarm(Y_j)
BPNN.biases = optimize_with_fish_swarm(Y_j)
    for epoch in range(max_epochs):
        output = BPNN.forward(input_data)
        error = calculate_MSE(output, expected_output)
        gradients = backpropagate(error)
BPNN.weights = BPNN.weights - η * gradients.weights
BPNN.biases = BPNN.biases - η * gradients.biases
```

Step 5: Extinction and Regeneration

```
remove_weakest_fish()
regenerate_strong_fish()
```

Step 6: Anomaly and Fault Detection

```
anomaly_score = BPNN.predict(test_data)
fault_score = BPNN.predict(test_data)
    if anomaly_score> threshold or fault_score> threshold:
trigger_early_warning()
```

Step 7: Return the optimized BPNN model for PS monitoring
Return BPNN model optimized using IFSO

# 4 Result and discussion

This section compares the result of the proposed method, an enhanced IFSO-BPNN framework, for AD and FD early warning in PS monitoring with existing methods. The evaluation was conducted using parameters such as accuracy (Acc), success rate (SR), misclassification instances (MI), error rate (ER), precision (Pre), recall (Rec), and F1 score (F1).

## 4.1 Experimental setup

The IFSO-BPNN technique is implemented on a machine equipped with an Intel i7 CPU, 16GB RAM, and a 512GB SSD. Python 3.9 is used for implementation, including libraries like NumPy, TensorFlow, Scikit-learn, and Matplotlib for processing and visualization. Table 2 displays the hyperparameters of the proposed method.

Table 2: Hyperparametric for proposed method

| **Hyperparameter** | **Range/Value** |
|---|---|
| BPNN Learning Rate (η) | 0.01 to 0.1 |
| Max Iterations (max_iter) | 100 to 1000 |
| Swarm Population Size | 50 to 200 |
| Max Step Size (max_s) | 0.1 to 1.0 |
| Min Step Size (min_s) | 0.01 to 0.1 |
| Learning Rate (η) for BPNN | 0.001 to 0.01 |
| Fitness Function | Error of BPNN model predictions |
| MSE Threshold for Convergence | 0.001 |
| Activation Function for BPNN | Mish, TanhExp, or ReLU |

## 4.2 Performance outcome

Figures 3 and 4 show the ROC curve and confusion matrix for anomaly detection and fault detection, respectively. The performance was evaluated based on the false positive rate, the true positive rate for the ROC curve, and the predicted and actual for the confusion matrix.

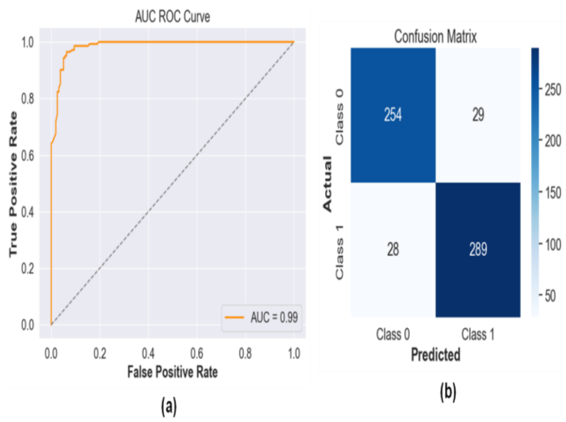


Figure 3: Anomaly detection (a) Roc curve, and (b) confusion matrix.

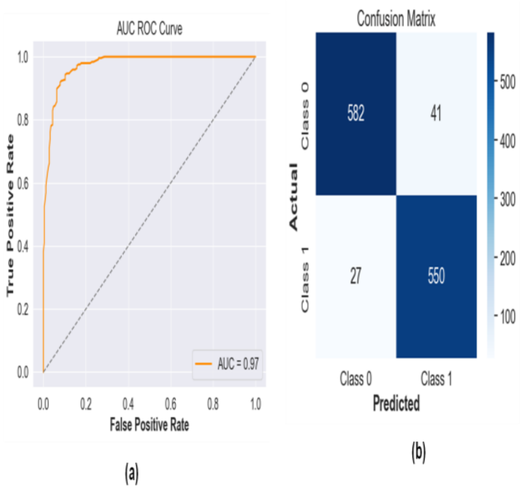


Figure 4: fault detection (a) Roc curve, and (b) confusion matrix.

## 4.3 Parameter explanation

Accuracy (Acc):Acc is defined as the ratio of accurately predicted occurrences (including true positives and true negatives) to total instances in a dataset, which measures the overall performance of PS monitoring and fault detection.Success rate (SR): The smart grid system is calculated as the proportion of accurately discovered faults and successful predictions to improveFD and maintenance accuracy.Misclassification instances (MI): The events occur when the model incorrectly identifies problems or normal conditions to demonstrate the possible flaws in identifying power defects.Error rate (ER):The fraction of misclassified cases, revealing the model's errors with an emphasis on decreasing mistakes in FD for PS.Precision (Pre) is the fraction of successfully diagnosed errors among all expected anomalies, demonstrating detection accuracy. Recall (Rec) measures the model's ability to detect all real abnormalities. F1 Score (F1) balances precision and recall. These metrics assess the IFSO-BPNN model's ability to accurately detect and monitor PS faults.

## 4.4 Comparison phase

The proposed method, IFSO-BPNN, is compared to the existing methods like Long Short-Term Memory (LSTM) [18] for FD, k-Nearest Neighbors (KNN), Decision tree classifier (DTC), and Random Forest (RF) [19] for ADand early warning in PS monitoring with evaluation metrics.
Table 3 and Figure 5 (a-b) display the comparison of metric values for the proposed method and existing methods to predict FD and FD in early warning of PS monitoring. The proposed IFSO-BPNN (98.5%) method achieves greater Acc than LSTM (91.21%).

Table 3: FD metrics values for proposed method.

| Metrics | LSTM [18] | IFSO-BPNN [Proposed] |
|---|---|---|
| Acc (%) | 91.21 | 98.5 |
| SR(%) | 92.42 | 96.85 |
| MI | 17 | 9 |
| ER (%) | 8.76 | 5.15 |

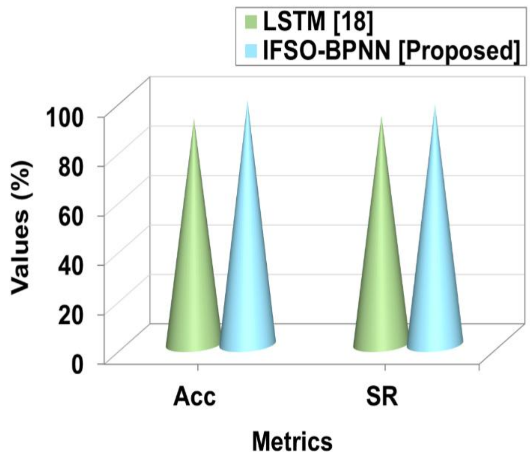


Figure 5(a): Acc and SR value for FD.

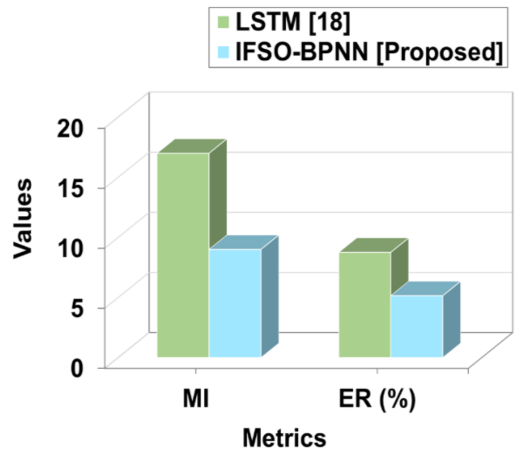


Figure 5(b): MI and ER FD value for proposed method.

Table 4 and Figure 6show the comparison of the proposed method and existing methods to evaluate the metric valuesused to predict ADand early warning of PS monitoring. The proposed IFSO-BPNN (0.9980) method achieves greater Acc than KNN (0.9729), DTC (0.9937) and RF (0.9976).

Table 4: Metrics values for proposed vs existing methods.

| **Metrics** | **KNN [19]** | **DTC [19]** | **RF [19]** | **IFSO-BPNN [Proposed]** |
|---|---|---|---|---|
| Pre | 0.9732 | 0.9937 | 0.9976 | 0.9978 |
| Rec | 0.9729 | 0.9937 | 0.9976 | 0.9977 |
| F1 | 0.9729 | 0.9937 | 0.9976 | 0.9979 |
| Acc | 0.9729 | 0.9937 | 0.9976 | 0.9980 |

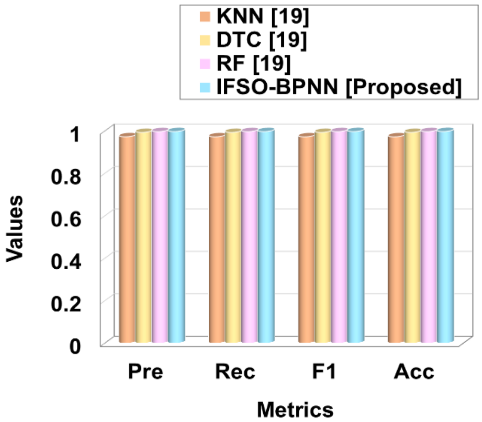


Figure 6: Evaluation metrics values for the proposed method.

In this research, both BPNN and IFSO-BPNN techniques were trained for FD and AD in PS. The numerical results of the ablations study for FD and AD in PS are displayed in Table 5, indicating that IFSO-BPNN performs better than BPNN.

Table 5: Outcome of ablation study

| **Method** | **AD Acc (%)** | **FD Acc (%)** |
|---|---|---|
| BPNN | 98.0 | 98.2 |
| IFSO-BPNN | 99.8 | 98.5 |

### 4.5 Discussion

The proposed IFSO-BPNN method achieves higher Acc, Pre, Rec, F1 and SR and significantly reduces MI and ER compared to existing methods like LSTM, KNN, DTCand RF. Existing models struggle with real-time adaptation and FD accuracy. The IFSO method overcomes these constraints by improving global search and optimizing BPNN parameters for improved performance. The connection helps electricity systems identify faults and provide early warnings. The key benefit is the substantial dependability and precision in predictive maintenance, which improves the robustness and efficiency of PS. Deploying the IFSO-BPNN model in smart grids provides real-time defect detection, such as detecting transformer overheating early on, averting blackouts, lowering maintenance costs, and enhancing energy distribution reliability across locations.

## 5 Conclusions

The improved early warning model, combining IFO with a BPNN (IFSO-BPNN), was presented to improve FD and predictive maintenance in smart power systems. The method aims to optimize neural network parameters for higher detection accuracy. The results demonstrated exceptional performance with FD accuracy (98.5%) and AD accuracy (0.9980) higher than existing methods. To address statistical validation, the IFSO-BPNN model has limited specificity, required more processing resources, and relied on precise parameter adjustment, which could leave an impact on real-time performance and generalizability across different power systems. The dataset's limited coverage of Beijing's local distribution stations, as well as a lack of sample size and class distribution information, limit its generalizability and model performance assessment. The future scope may extend the dataset to cover varied power systems, and providing precise details on sample size and class distribution would improve model resilience, generalization, and performance evaluation. Future research should focus on increasing specificity, testing in a variety of grid scenarios, and incorporating real-time adaptive processes to widen and improve the system's FD capabilities and use confidence intervals and standard deviations to demonstrate dependability. Future directions include statistical validation methods, such as confidence intervals and standard deviations, to support the reliability of results, providing clearer justification for performance metrics and model robustness. Future work will concentrate on providing thorough feature extraction, dimensionality reduction using ICA, and using correlation reduction methods for better analysis. Future research aims to enhance model performance and generalization by improving feature extraction, incorporating diverse data sources, and reducing dimensionality.

## References

[1] Wang G, Xie J, and Wang S (2023). Application of artificial intelligence in power system monitoring and fault diagnosis. *Energies*, 16(14), 5477. https://doi.org/10.3390/en16145477

[2] Chen Q, Li Q, Wu J, He J, Mao C, Li Z, and Yang B (2023). State monitoring and fault diagnosis of HVDC system via KNN algorithm with knowledge graph: a practical China power grid case. *Sustainability*, 15(4), 3717. https://doi.org/10.3390/su15043717

[3] He K, Wang T, Zhang F, and Jin X (2022). Anomaly detection and early warning via a novel multiblock-based method with applications to thermal power plants. *Measurement*, 193, 110979. https://doi.org/10.1016/j.measurement.2022.110979

[4] Stanković AM, Tomsovic KL, De Caro F, Braun M, Chow JH, Čukalevski N and Zhao S. (2022). Methods for analysis and quantification of power system resilience. *IEEE Transactions on Power Systems*, 38(5), 4774–4787. https://doi.org/10.1109/TPWRS.2022.3212688

[5] Bolbot V, Theotokatos G, Hamann R, Psarros G, and Boulougouris E (2021). Dynamic blackout probability monitoring system for cruise ship power plants. *Energies*, 14(20), 6598. https://doi.org/10.3390/en14206598

[6] Baba M, Nor NB, Sheikh MA, Baba AM, Irfan M, Glowacz A, and Kumar A(2021). Optimization of phasor measurement unit placement using several proposed case factors for power network monitoring. *Energies*, 14(18), 5596. https://doi.org/10.3390/en14185596

[7] Florkowski M (2021). Anomaly detection, trend evolution, and feature extraction in partial discharge patterns. *Energies*, 14(13), 3886. https://doi.org/10.3390/en14133886

[8] Du J, Wang X, and Zhang H (2025). Secure power management in wireless sensor networks for power monitoring using deep reinforcement learning. *Informatica*, 49(19). https://doi.org/10.31449/inf.v49i19.7125

[9] Qin G, Juan M, and Rui MH (2025). IoT-based intelligent power supply management using ensemble learning for seismic observation stations. *Informatica*, 49(8). https://doi.org/10.31449/inf.v49i8.6502

[10] Xiang L, Yang X, Hu A, Su H, and Wang P (2022). Condition monitoring and Anomaly detection of wind turbine based on cascaded and bidirectional deep learning networks. *Applied Energy*, 305, 117925. https://doi.org/10.1016/j.apenergy.2021.117925

[11] Chen H, Liu H, Chu X, Liu Q, and Xue D (2021). Anomaly detection and critical SCADA parameters identification for wind turbines based on LSTM-AE neural network. *Renewable Energy*, 172, 829–840. https://doi.org/10.1016/j.renene.2021.03.078

[12] Said AM, Yahyaoui A, and Abdellatif T (2021). Efficient Anomaly detection for smart hospital IoT systems. *Sensors*, 21(4), 1026. https://doi.org/10.3390/s21041026

[13] Zhang C, Hu D, and Yang T (2022). Anomaly detection and diagnosis for wind turbines using long short-term memory-based stacked denoising autoencoders and XGBoost. *Reliability Engineering & System Safety*, 222, 108445. https://doi.org/10.1016/j.ress.2022.108445

[14] Božiček A, Franc B, and Filipović-Grčić B (2022). Early warning weather hazard system for power system control. *Energies*, 15(6), 2085. https://doi.org/10.3390/en15062085

[15] He H, Sun F, Wang Z, Lin C, Zhang C, Xiong R, and Zhai L (2022). China's battery electric vehicles lead the world: achievements in technology system architecture and technological breakthroughs. *Green Energy and Intelligent Transportation*, 1(1), 100020. https://doi.org/10.1016/j.geits.2022.100020

[16] Bento ME (2022). Monitoring of the power system load margin based on a machine learning technique. *Electrical Engineering*, 104(1), 249–258. https://doi.org/10.1007/s00202-021-01274-w

[17] Adumene S, Islam R, Amin MT, Nitonye S, Yazdi M, and Johnson KT (2022). Advances in nuclear power system design and fault-based condition monitoring towards the safety of nuclear-powered ships. *Ocean Engineering*, 251, 111156. https://doi.org/10.1016/j.oceaneng.2022.111156

[18] Veerasamy V, Wahab NIA, Othman ML, Padmanaban S, Sekar K, Ramachandran R, and Islam MZ (2021). LSTM recurrent neural network classifier for high impedance fault detectionin solar PV integrated power system. *IEEE Access*, 9, 32672–32687. https://doi.org/10.1109/ACCESS.2021.3060800

[19] Mokhtari S, Abbaspour A, Yen KK, and Sargolzaei A (2021). A machine learning approach for anomaly detection in industrial control systems based on measurement data. *Electronics*, 10(4), 407.https://doi.org/10.3390/electronics10040407