processes

MDPI

Article

Interpretable Intelligent Fault Diagnosis for Heat Exchangers Based on SHAP and XGBoost

Haibo Xu [1], Shiwei Yan [2,3], Xiaogang Qin [1], Weizheng An [1] and Jinjiang Wang [3,4,*]

[1] CNOOC China Limited Beijing Research Center, Beijing 100028, China
[2] School of Mechanical and Transportation Engineering, China University of Petroleum, Beijing 102249, China
[3] Key Laboratory of Oil and Gas Production Equipment Quality Inspection and Health Diagnosis, State Administration for Market Regulation, Beijing 100088, China
[4] School of Safety and Ocean Engineering, China University of Petroleum, Beijing 102249, China
* Correspondence: jwang@cup.edu.cn

Abstract: Heat exchangers play an important role in offshore oil and gas production and offshore low-temperature waste heat power generation. Heat exchanger health management mainly relies on performance parameter monitoring; however, accurate fault mode identification remains a challenge and the diagnostic results are difficult to interpret. This paper proposed an interpretable intelligent diagnosis method for heat exchangers based on Shapley Additive exPlanation and eXtreme Gradient Boosting. Firstly, secondary parameters characterizing the health state of heat exchangers are constructed based on online monitored data such as pressure and temperature. Then, the multi-dimensional feature vector is constructed by integrating the online monitored data and secondary parameters. The XGBoost fault diagnosis model for leakage and scaling faults is established with the multi-dimensional feature vector as the input. To improve the model performance on diagnosis, a grid search optimization algorithm is adopted. After the fault diagnosis of a heat exchanger is completed, the SHAP method is introduced to analyze the contribution of features by quantifying their influence on the fault diagnosis results. Finally, the effectiveness of the proposed method is verified by the heat exchanger fault simulation experiment. The proposed method achieved a diagnosis accuracy of 99.79%, which has better performance than the traditional fault diagnosis method.

Keywords: heat exchanger; interpretable fault diagnosis; feature contribution analysis; XGBoost

check for updates

Academic Editor: Massimo Caruso

Received: 16 December 2024
Revised: 8 January 2025
Accepted: 11 January 2025
Published: 14 January 2025

Citation: Xu, H.; Yan, S.; Qin, X.; An, W.; Wang, J. Interpretable Intelligent Fault Diagnosis for Heat Exchangers Based on SHAP and XGBoost. *Processes* **2025**, *13*, 219. https://doi.org/10.3390/pr13010219

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

As key equipment for heat transfer, heat exchangers play an important role in offshore oil and gas production and low-temperature waste heat power generation, such as fluid cooling and heating of the working medium. Heat exchangers often work in harsh conditions like those with high temperature, pressure, and radiation, as well as mechanical load and corrosion, which can lead to failures such as leakage and fouling under long-term operation. Its performance and reliability are affected, and security accidents can even occur. According to statistics, there is a fouling phenomenon that exists in more than 90% of heat exchangers [1], and approximately 10% of corrosion failure in industrial systems results from internal media leakage in heat exchange equipment. Therefore, it is necessary to establish the fault diagnosis model of heat exchangers to accurately identify the health state and prevent or reduce the occurrence of sudden accidents.

The traditional fault diagnosis method of heat exchangers is mainly based on the mechanism model or expert knowledge. In the past, scholars have conducted a lot of

research on hydraulic [2], thermodynamic calculation [3,4], and other mechanisms [5]. For instance, Wallhaßer et al. [6] proposed a heat exchanger fouling evaluation method based on the pressure drop, heat transfer, and other indexes. Wang et al. [7] proposed a virtual sensing method for heat exchanger diagnosis by integrating failure mechanisms and inference analytics. Patil et al. [8] studied thermal and hydraulic fouling models and made indexes for fouling trend evaluation. The method based on this mechanism has good performance in fault diagnosis under a single working condition, but for complex and variable working conditions, the accuracy of the diagnosis declines due to modeling errors or parameter uncertainties.

To address the shortcomings of traditional mechanism-based approaches, intelligent fault diagnosis methods based on machine learning have been gradually applied to the diagnosis of heat exchangers in recent years [9–11]. With large amounts of data collected from sensors, machine learning models are able to identify patterns in the data and automatically classify fault types. Davoudi et al. [12] constructed a fouling factor based on the monitored data in the system and achieved the fouling degree evaluation using an artificial neural network. Hacene et al. [13] proposed a fuzzy model-based fault detection and diagnosis method for heat exchangers. Liang et al. [14] built a wavelet neural network based on improved particle swarm optimization for heat exchanger fouling fault assessment. Tian et al. [15] built a fault detection model for heat exchangers based on a dynamic mechanism model and the long short-term memory method, which has good performance in the early warning of internal leakage fault. Jallal et al. [16] proposed a fouling fault detection model for heat exchangers by integrating neural networks and principal component analysis. However, although data-driven intelligent diagnosis methods have achieved significant improvement in accuracy, most of these models are "black boxes". It is difficult for people to intuitively understand how the model extracts features from the input data and guides the model to make decisions. Shapley Additive exPlanation (SHAP) [17] is a model agnostic explanation method that can be flexibly adapted to any model and is highly adaptable. The advantage is that black-box models can be interpreted without sacrificing predictive effectiveness. As a reliable interpretation technique, SHAP has been widely used in the field of fault diagnosis in recent years.

Other common heat exchanger fault detection methods are based on electromagnetic probes [18], optical technology [19], acoustics [20], or water radioactivity analysis. However, the application of these methods is always offline and requires specialized detection equipment, limiting their use for real-time monitoring and fault detection.

In this paper, an interpretable intelligent diagnosis method based on SHAP and eXtreme Gradient Boosting (XGBoost) is proposed to solve the problems of inaccurate fault diagnosis and the unclear interpretation of diagnosis results in multiple fault modes of heat exchangers. The major works and contributions can be summarized as follows:

(1) In this proposed fault diagnosis framework, multi-dimensional feature vector construction, fault diagnosis, and model explanation for heat exchangers are integrated.
(2) For fault diagnosis, secondary parameters characterizing the health state of heat exchangers are constructed based on the online monitored data such as pressure and temperature. By integrating the online monitored data and secondary parameters, multi-dimensional feature vectors are obtained, and the XGBoost fault diagnosis model for leakage and fouling fault identification is established.
(3) The SHAP algorithm is introduced and integrated with XGBoost, which can analyze the feature contribution of the heat exchanger diagnosis model by quantifying their influence on the fault diagnosis results.

This paper is structured as follows. In Section 2, the theoretical foundation of XGBoost and SHAP is illustrated. In Section 3, the proposed interpretable intelligent diagnosis

method is described. In Section 4, a heat exchanger fault simulation experiment is carried out to illustrate the diagnosis performance of the proposed method. In Section 5, the conclusions are given.

2. Preliminary

2.1. XGBoost

In this paper, the XGBoost algorithm is used to model the fault mode recognition of heat exchanger leakage and scaling. XGBoost, first proposed by Chen et al. [21] in 2016, is an efficient algorithm that integrates multiple weak learners to improve the classification effect and improve model prediction by integrating multiple trees (usually decision trees). Its core idea is to use gradient descent optimization to build new trees in each iteration and gradually reduce the error of the current model.

As shown in Equation (1), XGBoost is an additive model composed of multiple base models, that is, by superimposing the predictions of multiple trees to generate the final predicted output $\hat{y}_i$, where f_k represents the prediction function of the Kth tree, x_i is the sample of the input, and K is the number of trees.

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i),\ f_k \in \varphi \tag{1}$$

The objective function of XGBoost consists of two parts, as shown in Equation (2): the error term $l(y_i, \hat{y}_i)$ and the regularization term $\Omega(f_k)$. The error term measures the errors between the predicted value and the true value, and the regularization term controls the complexity of the model. The formula is:

$$\begin{aligned} Obj &= \sum_{i=1}^{m} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \\ \Omega(f_k) &= \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \end{aligned} \tag{2}$$

In the formula, y_i is the sample to be predicted; γ is the regularization parameter of the number of leaf nodes, which is used to inhibit the continued splitting of nodes; T is the depth of the tree, which means the number of leaf nodes in the tree; and λ is the minimum sample weight of the leaf node to prevent the weight of the leaf node from being too large. w_j is the value of each leaf node.

During the training process, the following objectives will be minimized during round iteration:

$$Obj^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{3}$$

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta \sum_{k=1}^{K} f_k^t(x_i) \tag{4}$$

where the $\hat{y}_i^{(t)}$ and $\hat{y}_i^{(t-1)}$ are, respectively, the prediction value in the t and $t-1$ iteration, and η is the learning rate. In order to prevent overfitting, the algorithm reduces the influence of each tree by adding a penalty to the number of leaf nodes, limiting the growth of leaf nodes, and introducing a learning rate.

2.2. Shapley Additive exPlanation

Shapley Additive exPlanation (SHAP) is a model interpretation method that quantifies the contribution of each feature to model predictions based on Shapley values in cooperative game theory. The feature here is an attribute or variable used to describe a sample. Features

are usually the data input to the model. SHAP is a model agnostic interpretative approach that provides consistent and locally accurate interpretation.

The core of the SHAP treats each feature as a "player" contributing to the prediction and then calculates the marginal contribution of each feature across all possible combinations. Marginal contribution is the incremental change to the model output when a feature is added from one subset to another. SHAP calculates the average of the marginal contributions of all features to determine each feature's impact on the predicted outcome.

For the model prediction results, assuming the feature set is S, the Shapley value ϕ_i of feature i is defined as:

$$\phi_i = \sum_{S \subseteq M \backslash \{i\}} \frac{|S|!(|M| - |S| - 1)!}{|M|!} [v(S \cup \{i\}) - v(S)] \tag{5}$$

where M is the set of all features; $S \subseteq M \backslash \{i\}$ represents a subset of features that do not contain feature i; $v(S)$ represents the model output that considers only the features in S. $|S|$ and $|M|$ are the number of ways that features in S and M are arranged. SHAP calculates the Shapley value for each feature and draws an interpretation map, ranking the contribution of the feature by size, which helps visualize the significance and positive or negative impact of the feature.

2.3. Model Evaluation Index

In order to evaluate the effectiveness of the fault diagnosis model, accuracy, recall, and F1-score index are introduced, which are defined as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{6}$$

$$\text{F1} - \text{score} = \frac{2 \cdot \text{precison} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{7}$$

True Positive (TP) refers to the case number where both the actual and predicted values are positive. False Positive (FP) represents the case number where the actual value is negative but the prediction is positive. False Negative (FN) denotes the case number where the actual value is positive but the prediction is negative. True Negative (TN) denotes the case number where both the actual and predicted values are negative. Precision and recall are calculated by using Equations (8) and (9).

$$\text{precision} = \frac{TP}{TP + FP} \tag{8}$$

$$\text{recall} = \frac{TP}{TP + FN} \tag{9}$$

When evaluating the performance of fault diagnosis model, the higher the value of accuracy, recall, and F1-score, the better the performance of the model.

3. Interpretable Intelligent Fault Diagnosis Based on SHAP and XGBoost

This section presents an interpretable intelligent diagnosis method for heat exchangers based on SHAP and XGBoost. As shown in Figure 1, this method combines data acquisition and health index construction, intelligent diagnosis, and model interpretation to realize intelligent fault diagnosis and decision support of a heat exchanger. Firstly, through multi-source health information perception, key monitored data are collected and secondary indicators reflecting equipment health status are constructed. Then, the XGBoost model is constructed by integrating the online monitored data and the secondary parameters to

realize the intelligent diagnosis of the leakage and scaling faults. Finally, the contribution of features was evaluated quantitatively by SHAP interpretation analysis, which made the diagnosis results more reliable and interpretable.

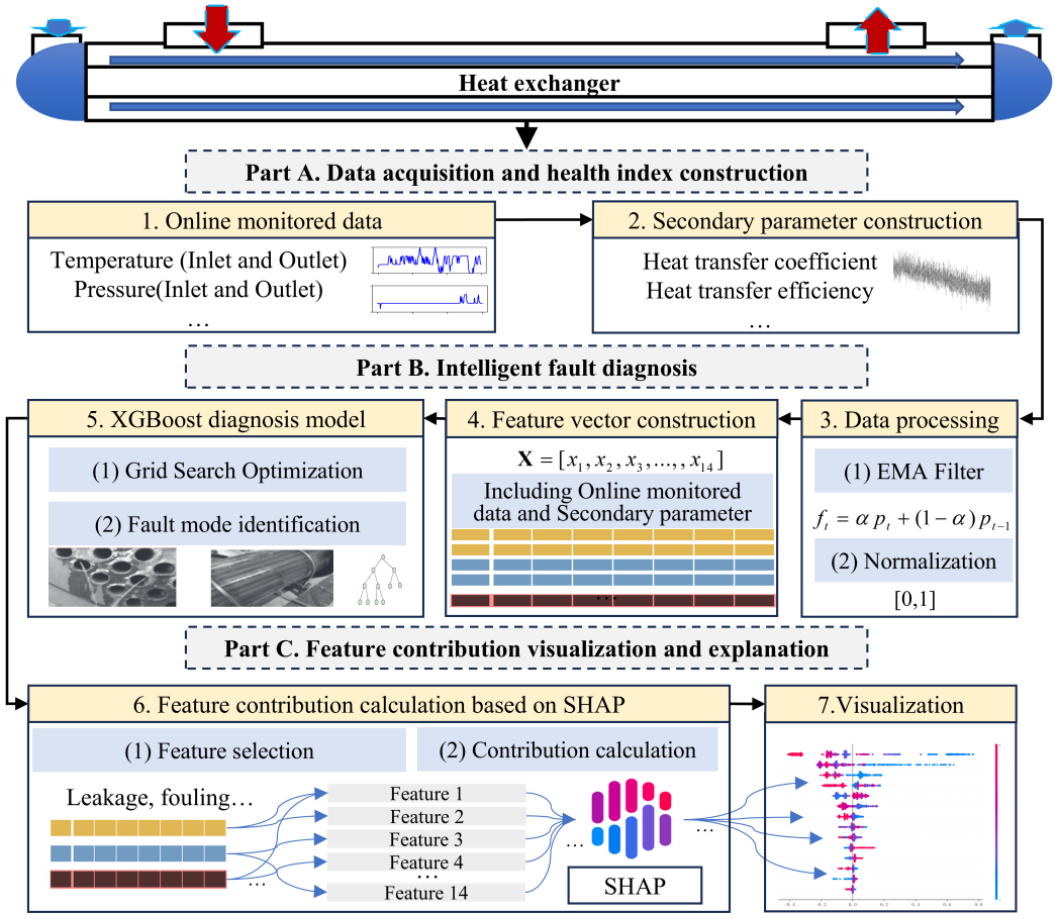


Figure 1. Framework of the proposed interpretable intelligent fault diagnosis method.

3.1. Data Acquisition and Data Processing

First of all, the temperature, pressure, pressure drop, and other online monitored data of the heat exchanger are continuously obtained through online monitoring. As shown in Table 1, based on the data of online monitored data, secondary parameters such as heat transfer coefficient and average temperature differential are constructed to further characterize the health state of the heat exchanger. These secondary indexes can not only reflect the heat transfer performance of the heat exchanger but also reveal its fault characteristics, providing more reliable input information for subsequent intelligent diagnosis. For the meaning of some intermediate parameters in Table 1, G_h and G_c are the mass flow rate of cold and hot fluids, respectively, and A is the heat exchange area of the heat exchanger. $t_1' - t_2'$ is the maximum temperature difference possible for the fluid. $(t - t'')_{\max}$ is the greater value of the actual temperature difference between cold and hot fluids. Φ is the actual heat transfer. $C_{\min}$ is the smaller value of the heat capacity flow of cold and hot fluids.

Table 1. Online monitored data and secondary parameters construction.

Parameters	Name	Abbreviation and Formula
Online monitored data	Hot side inlet temperature	T_{hi}
	Hot side outlet temperature	T_{ho}
	Cold side inlet temperature	T_{ci}
	Cold side outlet temperature	T_{co}
	Hot side inlet pressure	P_{hi}
	Cold side inlet pressure	P_{ci}
	Hot side pressure drop	$H_pressure_drop$
	Cold side pressure drop	$C_pressure_drop$
Secondary parameters	Average temperature differential	$\Delta T_m = \frac{(T_{hi}-T_{ci})-(T_{ho}-T_{co})}{\lg(T_{hi}-T_{ci})-(T_{ho}-T_{co})}$
	Heat transfer coefficient	$K = \Phi/(A\Delta T_m)$
	Entropy production value	$Ent_pro_value = \ln(T_{ho}/T_{hi})/W + \ln(T_{co}/T_{ci})$
	Heat transfer effectiveness	$HT_effectiveness = (t - t'')_{\max}/(t_1'' - t_2')$
	Flow power consumption	$Fpower_consumption = \left[G_h(P_{hi} - P_{ho}) + G_c(P_{ci} - P_{co})\right]$
	Number of heat transfer units	$NTU = (KA)/C_{\min}$

As shown in Equation (10), the exponential moving average (EMA) is used to reduce the noise of online monitored data and secondary parameters so as to eliminate the parameter fluctuations caused by external interference over time.

$$f_t = \alpha p_t + (1 - \alpha) p_{t-1} \tag{10}$$

where α represents the smoothing parameter whose value ranges from 0 to 1. p_t is the data to be processed; f_t indicates the data after EMA processing. When processing time series data, EMA can process real-time monitoring data and historical data in a unified manner. If the α value is larger, the real-time data have a larger weight in the EMA value of the current moment, whereas the historical data have a larger weight in the EMA value of the current moment. In practical applications, α is adjusted based on the data. Dimensional differences of features are eliminated through data normalization processing to improve data quality. A total of 14 features x_j, including primary monitoring parameters and secondary indicators, were spliced after pretreatment to construct a multidimensional feature vector $\mathbf{X}$, as shown in Equation (11).

$$\mathbf{X} = [x_1, x_2, x_3, \ldots, x_{14}] \tag{11}$$

3.2. XGBoost Diagnosis Model

As shown in Equation (12), the historical data of the device is divided into different health states to obtain dataset D for training the fault diagnosis model, where $\mathbf{X}_i$ is the input sample of the intelligent diagnosis model and y_i is the fault type label corresponding to the sample.

$$D = \{\mathbf{X}_i, y_i\}, i = 1, 2, 3 \ldots\ldots \tag{12}$$

XGBoost is used to model the fault mode recognition of heat exchanger leakage and scaling. In the process of model construction, the hyperparameter selection of the intelligent diagnosis model XGBoost is optimized by the grid search optimization algorithm, that is, by building a regular grid in the specified parameter space, traversing all possible hyperparameter combinations, training the model, evaluating the performance, and, finally, selecting the best parameter combination. As shown in Equation (13), the overall objective of the model parameter optimization process can be expressed as:

$$\theta^* = \arg\min_{\theta \in \Theta} (1 - \frac{1}{N} \sum_{i=1}^{N} \text{accuracy}_i) \tag{13}$$

where θ is the hyperparameter set; Θ is the hyperparameter search space; accuracy illustrated in Equation (6) is the performance evaluation function; N is the validation set size.

The goal of model optimization is to find the optimal parameter combination. Accuracy is used as a performance metric in this research. Each group of hyperparameter combinations is evaluated by cross-validation of the average accuracy on the validation set. The combination with the highest accuracy is finally selected as the optimization result, and the diagnostic model is constructed based on the optimal model parameters.

As shown in Equation (14), after model training is completed, the diagnosis output of XGBoost is an additive model composed of multiple base models. The fault-predicted label $\hat{y}_i$ is generated by superposing the predictions of multiple trees, where K is the number of trees.

$$\hat{y}_i = \sum_{k=1}^{K} f_k(\mathbf{X}_i), f_k \in \varphi \tag{14}$$

3.3. Feature Contribution Visualization and Explanation

In order to quantify the contribution of each feature to the fault diagnosis result during the diagnosis process, the SHAP model is introduced. The diagnosis result of the XGBoost is composed of multiple decision tree base functions f_k, and the Shapley value is calculated separately in each tree and finally accumulated. By calculating the average marginal contribution of all features, the influence of each feature on the predicted result is obtained. For the prediction result, the feature contribution based on the Shapley value is shown in Equation (15):

$$\text{Contribution}_{i,p} = \sum_{p} \frac{\text{weight}(p)}{\text{path}_{\text{count}}(p)} \cdot \Delta f_k(p, i) \tag{15}$$

where the $\text{contribution}_{i,p}$ is the contribution of feature i for the path p in the XGBoost tree model; $\text{weight}(p)$ is the weight of the path p, representing the probability of reaching the leaf node along the path; $\text{path}_{\text{count}}(p)$ is the number of feature combinations in path p; $\Delta f_t(p, i)$ represents the prediction increment of the feature i for the tree k on the path p.

By weighted summation of all path contributions, the Shapley value ϕ_i of features in the entire model can be obtained, as shown in Equation (16):

$$\phi_i = \sum_{t=1}^{T} \sum_{p} \text{Contribution}_{i,p} \tag{16}$$

By calculating the Shapley value of each feature and drawing the interpretation map, the contribution of the feature is ranked. And the positive or negative impact of the feature on the model are visualized.

4. Experiments

4.1. Experiment Setup

A fault simulation test bench was constructed to evaluate the proposed method, as illustrated in Figure 2. The test bench consists of three circuits: an extrinsic cycle circuit of the cold side, an intrinsic cycle circuit of the test, and an extrinsic cycle circuit of the hot side. A shell-and-tube heat exchanger is in the test circuit with internal leakage, external leakage, and fouling fault being simulated. The fluid with higher temperatures and pressures is in the tube side. During the experiment, temperature, pressure, and flow are monitored by using PT100 platinum thermal resistance sensors, diffused silicon pressure transmitters and turbine flowmeters. During the experiment, Siemens S7-1200 PLC is used for online data collection at the measurement points, gathering data on inlet/outlet temperatures, pressures, and flow rates. The sampling period of the simulation experiment is 10 s. Three kinds of internal leakage, external leakage, and fouling fault tests are carried out.

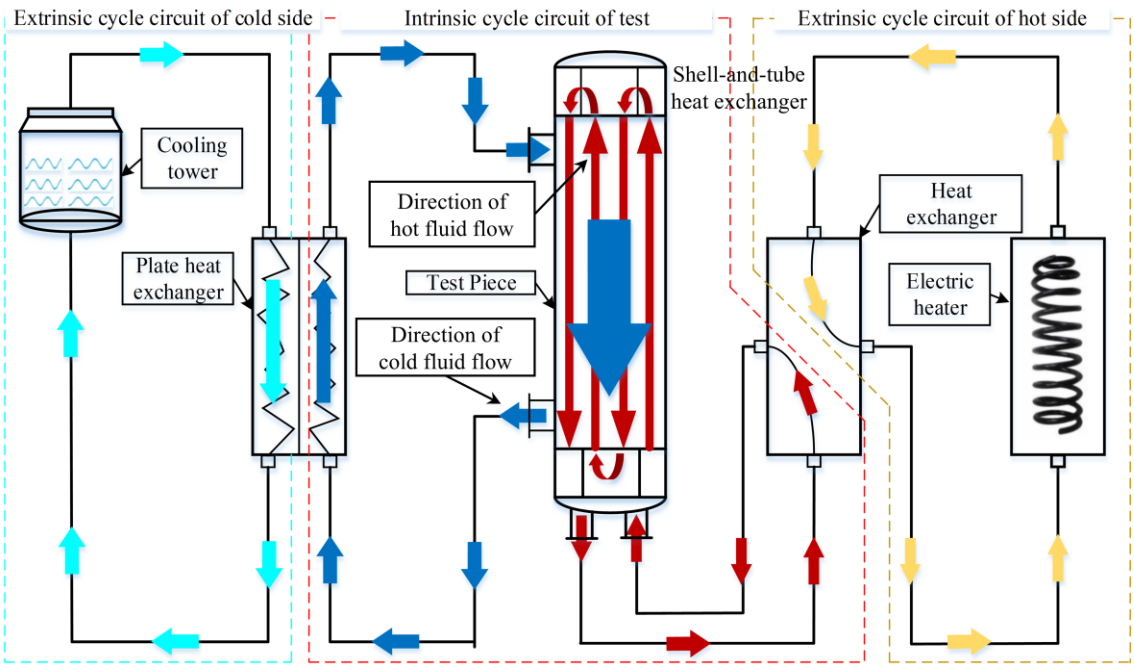


Figure 2. Test bench diagram.

The experiment set for leakage and fouling fault is shown in Table 2. For the fouling test of the heat exchanger, pure water and $CaSO_4$ solution are, respectively, in the tube side and shell side. For leakage faults, threaded holes are machined as leakage channels on the heat exchanger tube plate and tube. During the fouling fault simulation experiment, data are collected for more than 40,000 s with a sampling period of 10 s, and the leakage fault simulation experiment continuously lasts for more than 10,000 s with the same sampling period.

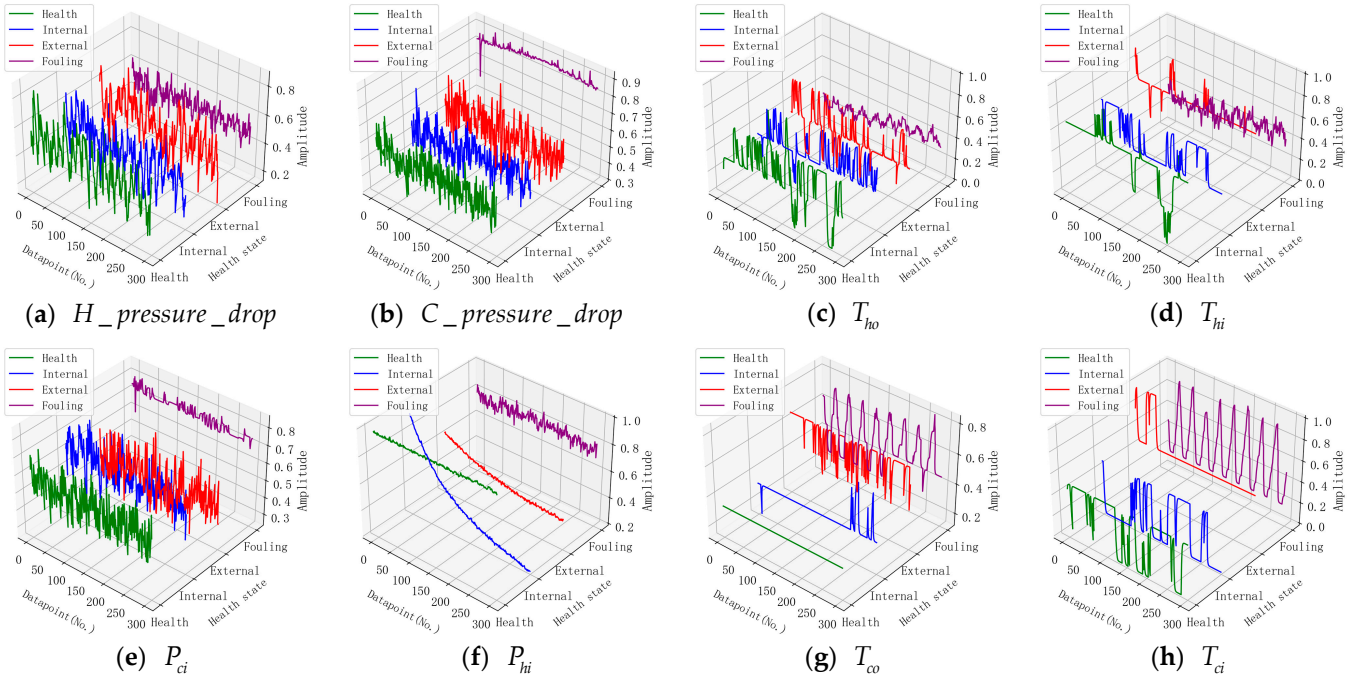
Table 2. Experiment set for leakage and fouling fault.

Parameters	Health State	Internal Leakage	External Leakage	Fouling
Hot/cold side inlet temperature (°C)	55/35	55/35	55/35	65/31
Hot/cold side flow rate (m^3/h)	25/20	25/20	25/20	0.72/0.44
Internal medium leakage flow (m^3/h)	\	0.15	\	\
External leakage flow of flange (m^3/h)	\	\	0.12	\
Fouling thickness (mm)	\	\	\	0.3

4.2. Data Analysis

The data acquired under different fault modes are shown in Figure 3. Eight types of original online monitored data such as inlet and outlet temperature, inlet and outlet pressure, and pressure drop on the hot/cold side are presented, respectively. For internal leakage, the hot side medium mixes with the cold side medium for direct heat transfer, causing an abnormal rise in the cold side outlet temperature. Whether it is an internal leakage or external leakage, when the leakage occurs, the flow rate of the medium is reduced, and the total flow area of the medium is increased. The flow channel resistance is reduced and the inlet pressure is reduced under the premise of providing the same power to the medium. The main influence on fouling lies in heat transfer. When fouling occurs, both sides of the wall need to resist greater thermal resistance under the same temperature difference, and the temperature difference will also be larger. Affected by variations in working conditions, the data amplitude of the outlet temperature and the inlet temperature fluctuate to a certain extent. The following multidimensional raw data

for visual evaluation, health, internal leakage, external leakage, and fouling failure modes are difficult to distinguish directly.



(**a**) $H_pressure_drop$ (**b**) $C_pressure_drop$ (**c**) T_{ho} (**d**) T_{hi}

(**e**) P_{ci} (**f**) P_{hi} (**g**) T_{co} (**h**) T_{ci}

Figure 3. Normalized online monitored data for four typical failure modes.

In this part, the parameter α of the filter EMA is adjusted according to the actual data. When the α value is close to 0.6, the abnormal fluctuation of the data is greatly reduced, and the dynamic characteristics of the monitored data with the change of the running state are effectively retained. When α exceeds 0.6, the dynamic characteristics of the monitored data are weakened. Therefore, the EMA smoothing parameter α is taken as 0.6 in this experiment. The normalized smoothed feature vector consisting of online monitored data is visualized by the *t*-SNE algorithm first. The *t*-SNE analysis results in Figure 4 show the feature distribution under different health conditions, respectively. Only by temperature-related parameters, in addition to fouling faults, a large number of samples of the three other fault modes of a healthy state, internal leakage, and external leakage, are aliased together. It is difficult to distinguish effectively. When the pressure-related parameters are used for visualization, the samples are roughly distributed in different areas, and the four health states can be effectively distinguished. However, a small number of samples overlap between the health state and the internal leakage, and a large number of samples overlap between the internal leakage and the external leakage. The sample discrimination effect is not ideal. When all the online monitored data, including temperature and pressure, are used for *t*-SNE analysis, most samples can be effectively distinguished. But there is still a small amount of sample overlap. The distribution characteristics indicate the difficulty of fault classification between the healthy state, leakage state, and fouling state. More effective diagnosis models need to be established to improve the accuracy of classification.

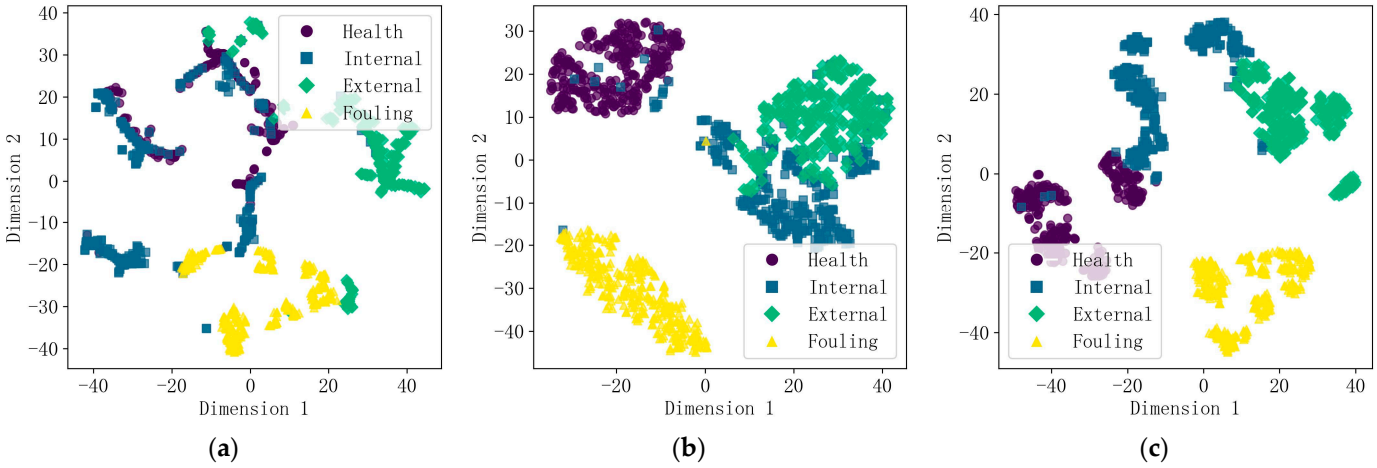


Figure 4. *t*-SNE visualization analysis for raw data: (**a**) temperature-related parameters; (**b**) pressure-related parameters; (**c**) all monitored parameters.

4.3. Diagnosis Results

First, for the training of the XGBoost diagnostic model, the feature vector is constructed based on the online monitored data obtained from the above test bench and constructed secondary parameters. Then, 300 samples are selected from different health states (including healthy state, internal leakage, external leakage, and fouling). A total of 180 samples of each health state were selected for the training of the XGBoost model, and the remaining 120 samples were left for the accuracy test of the diagnostic model. In order to have a better model performance on fault diagnosis, a grid search optimization method is adopted for hyperparameter selection of XGBoost model. The grid search optimization method instantiates the model through each combination in the list of hyperparameter combinations and returns the optimal parameters to the model object as the final diagnostic model parameters. The optimization results of the model parameters are shown in Table 3.

Table 3. The optimized XGBoost model parameters.

Model Parameter	Parameter Search Range	Parameter Search Result
Learning Rate	0.01, 0.1, 0.2	0.2
Maximum Depth	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	6
Column Subsample by Tree	0.2, 0.4, 0.6, 0.8, 1.0	0.4
Subsample	0.2, 0.4, 0.6, 0.8, 1.0	0.6

Figure 5 presents the curve of the loss function in the training and testing process of the model. It can be observed from the figure that with the increase of the number of iterations, the loss of the training set and the testing set gradually decrease and gradually become stable, and finally converge to a certain stable value, which indicates that the training effect of the model is ideal. At the same time, after the loss becomes stable, the loss values of the test set and the training set are close, which indicates that the model has good generalization ability on the test set.

Figure 6 shows the confusion matrix for fault diagnosis based on the XGBoost model. The matrix shows the classification results among healthy state, internal leakage, external leakage, and fouling faults, which are, respectively, labeled 0, 1, 2, and 3. Each fault category contains 120 samples. The XGBoost model achieved a 99.79% diagnosis accuracy, where only one sample was misclassified. It shows that the XGBoost model can identify various kinds of faults well and has high classification accuracy.

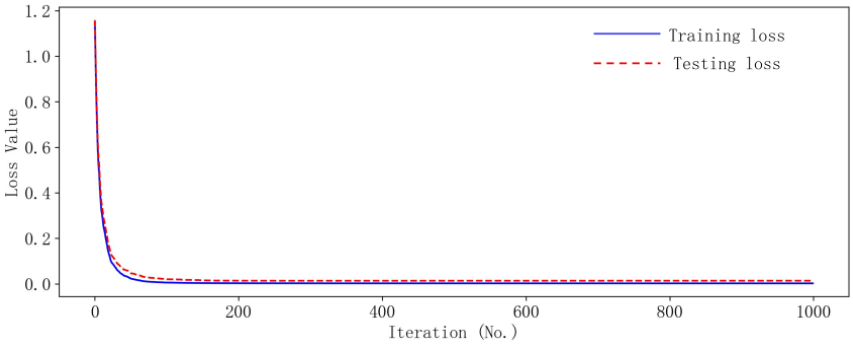


Figure 5. Loss function curve visualization.

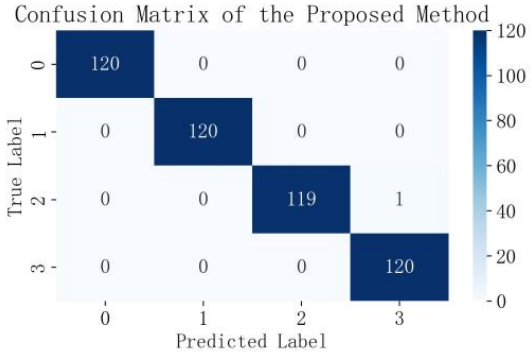
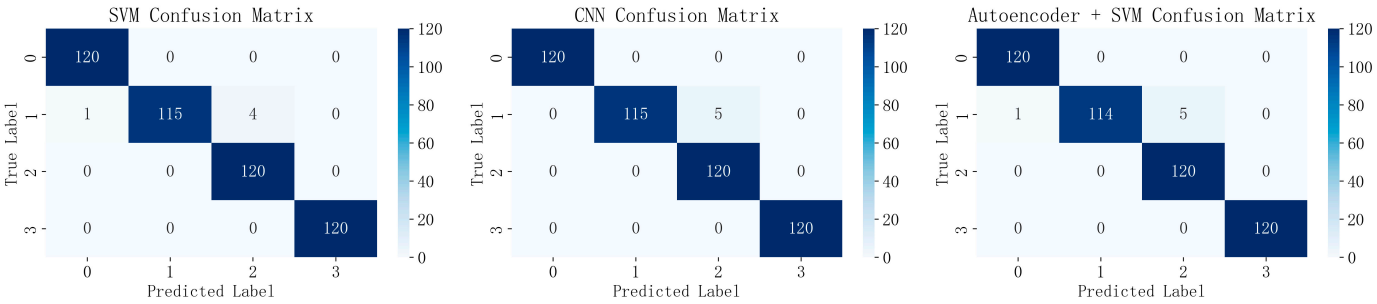


Figure 6. Fault diagnosis confusion matrix for XGBoost.

In order to further analyze the effectiveness of the proposed method, a diagnosis model based on Support Vector Machine (SVM), Convolutional Neural Network (CNN), and AutoEncoder + SVM were constructed to be compared and analyzed. The diagnosis confusion matrix is illustrated in Figure 7. It is clear that there are at least five samples misclassified based on the above three diagnosis models. The comparison results of the four models for accuracy, recall, and F1-score are shown in Figure 8 and Table 4. The XGBoost model shows the best fault diagnosis performance in all three model evaluation indexes with an accuracy, recall, and F1-score of almost 100%, indicating that it is highly efficient at dealing with the diagnosis problem. The accuracy of the CNN model is 99.00%, which is better than the SVM and AutoEncoder + SVM models but still slightly lower than that of XGBoost. But it also shows that the CNN has good performance on time series feature extraction and diagnosis problems. In contrast, the SVM and AutoEncoder + SVM models have slightly poorer performance with 98.99% and 98.79% accuracy, respectively. AutoEncoder + SVM is a two-stage method. AutoEncoder is first used to extract features, and then SVM is used to classify features. Theoretically, the feature extraction capability of deep learning and the classification capability of SVM should be taken into account. But this method performs the worst. It is possibly because the feature extraction is insufficient for low-dimension input data, and the features of AutoEncoder cannot fully reflect the fault characteristics with limited training data.

These comparison results further validate the superior performance of XGBoost model in this fault diagnosis task. It has higher reliability and accuracy when dealing with multi-fault diagnosis problems in this study.



(**a**) SVM (**b**) CNN (**c**) AutoEncoder + SVM

Figure 7. Confusion matrix comparison of diagnostic results by using different models.

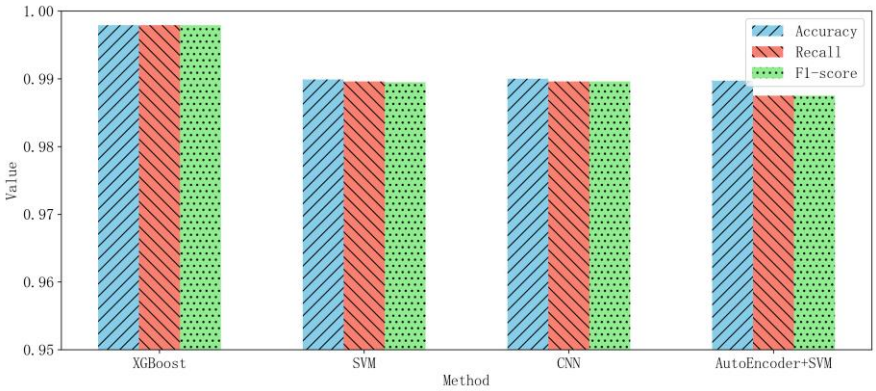


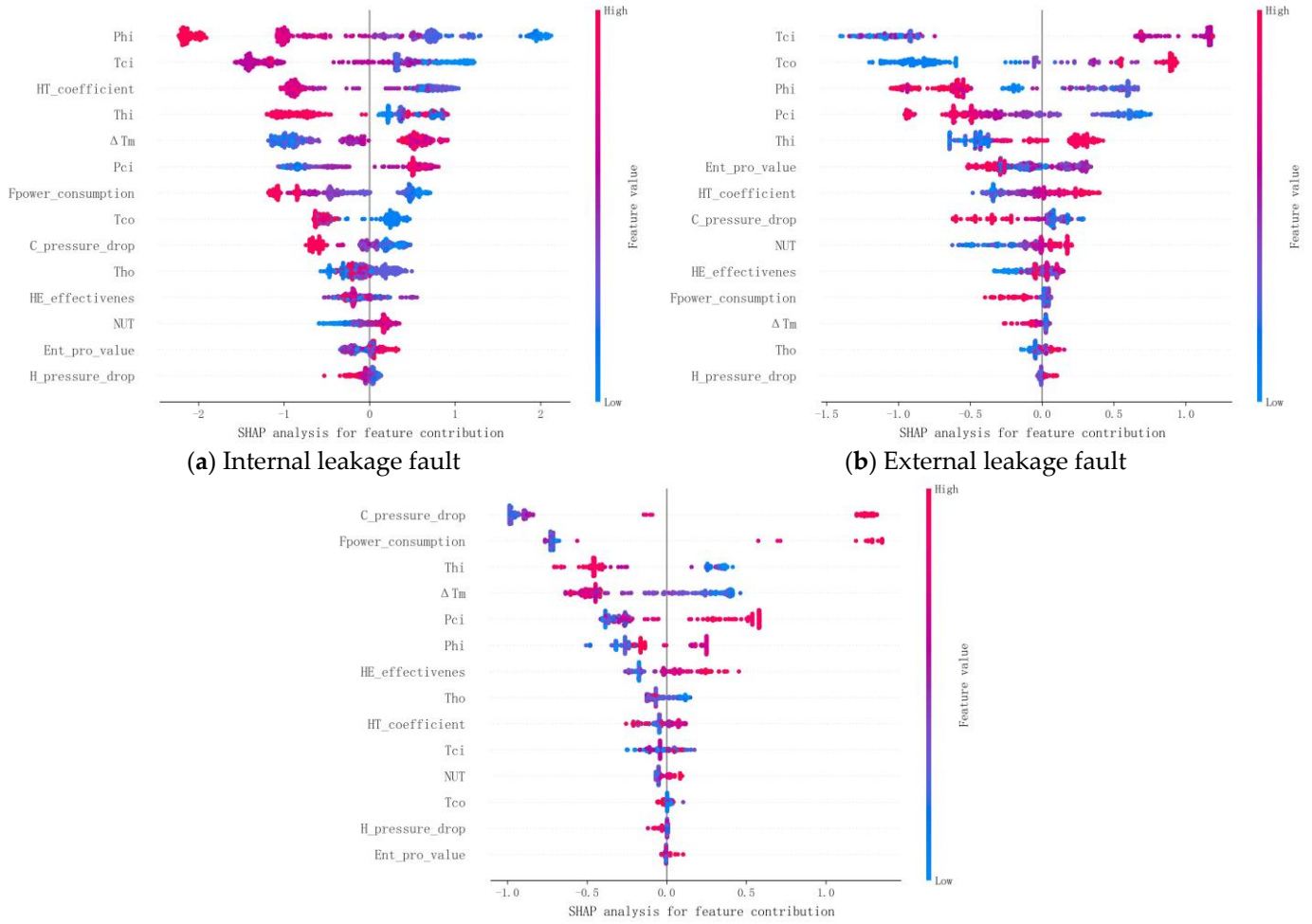
Figure 8. Comparison of model performance.

Table 4. Fault diagnosis result analysis based on different models.

Method	Accuracy	Recall	F1-Score
XGBoost	99.79%	99.79%	99.79%
SVM	98.99%	98.96%	98.95%
CNN	99.00%	98.96%	98.96%
AutoEncoder + SVM	98.79%	98.75%	98.75%

Figure 9 shows the visual results of SHAP values to explain the feature contribution to the XGBoost model in leakage and fouling fault diagnosis. The SHAP values represent the degree to which each feature contributes to the model prediction, where a positive value indicates that the feature pushes the prediction in the "fault" direction and a negative value pushes the prediction in the "normal" direction. The colors in the graph from red to blue indicate the change in feature values from high to low.

For internal leakage of the heat exchanger, the cold and hot fluids often mix directly, which will change the fluid flow characteristics of the hot side, especially the pressure distribution. As illustrated in Figure 9a, from the point of view of the mechanism, the hot side inlet pressure P_{hi} changes significantly due to the flow disturbance, which contributes most to internal leakage diagnosis. As the cold side fluid is mixed with the high-temperature fluid leaking from the hot side, the cold side inlet temperature T_{ci} is disturbed, which is reflected in temperature rise or fluctuation. For the cold side outlet temperature T_{co}, the cold side temperature increases after the mixture of cold and hot fluids, which directly reflects the interference of internal leakage on the heat transfer effect. After the direct mixture of cold and hot fluids, the heat transfer driving force (temperature difference) of the heat exchanger decreases, the heat transfer performance deteriorates, and its effect on the average temperature differential ΔT is significant.



(**a**) Internal leakage fault (**b**) External leakage fault

(**c**) Fouling fault

Figure 9. Feature contribution visualization based on SHAP.

External leakage is usually caused by the failure of the shell, joint, or seal of the heat exchanger, resulting in the leakage of the medium into the external environment. The direct consequence is a decrease in fluid flow, a significant change in pressure, and a decrease in heat transfer performance. The direct influence of external leakage on the cold side fluid is reflected in the decrease in flow rate, which leads to the change of temperature distribution on the cold side. Figure 9b indicates that the most important features are the cold side inlet temperature T_{ci}, cold side outlet temperature T_{co}, hot side inlet pressure P_{hi}, and cold side inlet pressure P_{ci}. T_{ci} and T_{co} play critical roles as external leakage leads to changes in cold side temperature distribution due to reduced flow rate. The fluctuation of inlet temperature is the initial manifestation of the leakage effect. External leakage will directly lead to a decrease in the flow rate of the cold side fluid, accompanied by a significant decrease in the cold side pressure, which is a direct characteristic of external leakage.

Fouling failures are usually caused by dirt deposits on the inner tube walls of heat exchangers. In Figure 9c, the most important features are the cold side pressure drop *C_pressure_drop*, flow power consumption *Fpower_consumption*, and hot side inlet temperature T_{hi}. This indicates that fouling primarily impacts fluid flow resistance and heat transfer performance. The fouling causes the flow channel to narrow and the fluid flow resistance to increase, which increases the flow power consumption and leads to the changes in pressure drop. SHAP values highlight *Fpower_consumption* (flow power consumption) as a critical feature, with higher values directly indicating fouling. Fouling affects the heat

transfer effectiveness *HT_effectiveness*, resulting in abnormal changes in temperature-related parameters such as ΔT_m. The reason for the higher contribution of cold side pressure drop *C_pressure_drop* in fouling diagnosis may also be that the working conditions are different from the leakage working conditions, which leads to a high influence in the model.

As shown in Figure 10, SHAP was further used to analyze the average contribution degree of online monitored data and secondary parameters under four health states. In summary, secondary parameters such as heat transfer coefficient, average temperature differential, flow power consumption, and others rank high in the contribution degree of fault diagnosis, indicating that the constructed secondary parameters have good separability in fault mode diagnosis. At the same time, the pressure parameter plays an important role in this study, and its average contribution degree ranks first. Through the analysis of the original data, it can be seen that when the internal leakage and external leakage faults occur, the change of the inlet pressure has a significant downward trend compared with other indicators, which also proves the rationality of the SHAP analysis results.

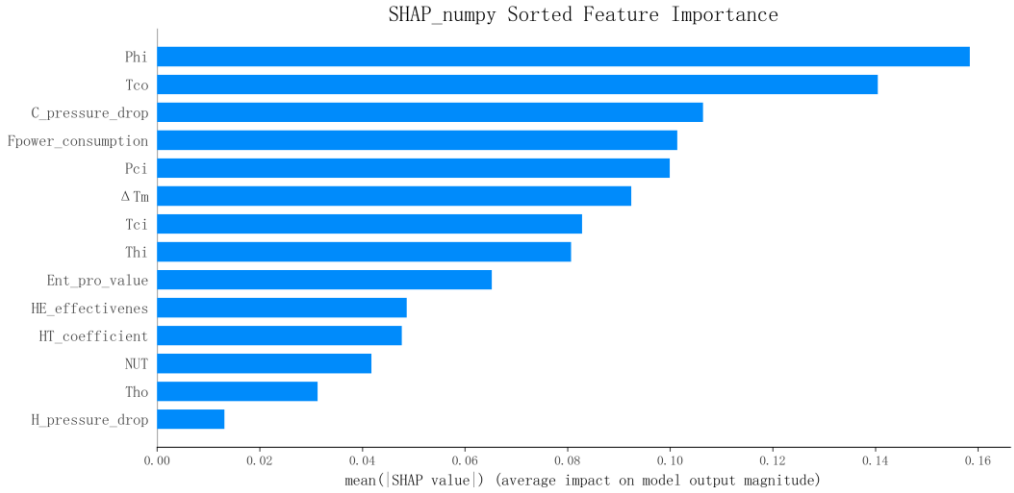


Figure 10. The average contribution of features on diagnosis model.

SHAP analysis reveals key features that have the most significant influence on the diagnosis model for leakage and fouling faults. The main influence of internal leakage on the condition monitoring parameters is the pressure fluctuation and temperature field redistribution caused by fluid mixture. The main influence of external leakage is the decrease of fluid flow and the pressure and temperature changes caused by it. The model also captures the feature's changes in the field of thermal performance and fluid resistance in fouling faults. These characteristics are clearly reflected in SHAP analysis, which can provide a basis for further optimization of the model and formulation of fault prevention strategies.

5. Conclusions

In order to solve the problem that the fault diagnosis for heat exchangers is not accurate and the diagnosis result is difficult to explain, this paper proposed an intelligent fault diagnosis method for heat exchangers based on SHAP and XGBoost. The main conclusions are as follows:

(1) The XGBoost intelligent fault diagnosis method integrating online monitored data and secondary parameters realizes an accurate identification of internal leakage, external leakage, and fouling faults. Compared with traditional SVM, CNN, and other methods, the diagnostic accuracy is improved by about 1%, showing better diagnostic performance.

(2) The XGboost combined with the SHAP model was used to analyze the sensitivity of features in different fault modes. The core influence of internal leakage is on pressure fluctuation and temperature field redistribution caused by fluid mixing. The main influence of external leakage is on fluid flow reduction and pressure and temperature changes caused by it. The fouling failure has a significant impact on thermal performance and fluid resistance. It has a guiding significance for field equipment maintenance personnel to deeply understand the cause of faults and provides strong support for the health management of heat exchangers.

For scalability and limitation, due to experimental constraints, this paper focuses solely on the shell-and-tube heat exchanger. The constructed features are based on heat transfer and fluid mechanics. For various kinds of heat exchangers, the feature calculation formula is slightly different, which needs to be adjusted for further application. For future work, the SHAP model can be used to select features of heat exchangers and optimize the construction of health indexes in order to reduce information redundancy and improve real-time diagnosis while ensuring the accuracy of fault diagnosis.

Author Contributions: Conceptualization, H.X. and J.W.; data curation, X.Q.; formal analysis, S.Y.; investigation, S.Y. and H.X.; methodology, X.Q. and S.Y.; project administration, J.W.; resources, J.W.; software, W.A; supervision, J.W.; validation, W.A. and H.X.; visualization, W.A.; writing—original draft, H.X.; writing—review and editing, X.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: Authors Haibo Xu, Xiaogang Qin and Weizheng An were employed by the CNOOC China Limited Beijing Research Center. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Steinhagen, R.; Müller-steinhagen, H.; Maani, K. Problems and Costs Due to Heat Exchanger Fouling in New Zealand Industries. *Heat Transf. Eng.* **1993**, *14*, 19–30. [CrossRef]
2. Fguiri, A.; Marvillet, C.; Jeday, M.R. Estimation of Fouling Resistance in a Phosphoric Acid/Steam Heat Exchanger Using Inverse Method. *Appl. Therm. Eng.* **2021**, *192*, 116935. [CrossRef]
3. Jonsson, G.R.; Lalot, S.; Palsson, O.P.; Desmet, B. Use of Extended Kalman Filtering in Detecting Fouling in Heat Exchangers. *Int. J. Heat Mass Transf.* **2007**, *50*, 2643–2655. [CrossRef]
4. Delrot, S.; Guerra, T.M.; Dambrine, M.; Delmotte, F. Fouling Detection in a Heat Exchanger by Observer of Takagi–Sugeno Type for Systems with Unknown Polynomial Inputs. *Eng. Appl. Artif. Intell.* **2012**, *25*, 1558–1566. [CrossRef]
5. Delmotte, F.; Dambrine, M.; Delrot, S.; Lalot, S. Fouling Detection in a Heat Exchanger: A Polynomial Fuzzy Observer Approach. *Control Eng. Pract.* **2013**, *21*, 1386–1395. [CrossRef]
6. Wallhäußer, E.; Hussein, M.A.; Becker, T. Detection Methods of Fouling in Heat Exchangers in the Food Industry. *Food Control* **2012**, *27*, 1–10. [CrossRef]
7. Wang, J.; Sun, J.; Ge, W.; Zhang, F.; Gao, R.X. Virtual Sensing for Online Fault Diagnosis of Heat Exchangers. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 9508708. [CrossRef]
8. Patil, P.D.; Kozminski, M.; Peterson, J.; Kumar, S. Fouling Diagnosis of Pennsylvania Grade Crude Blended with Opportunity Crude Oils in a Refinery Crude Unit's Hot Heat Exchanger Train. *Ind. Eng. Chem. Res.* **2019**, *58*, 17918–17927. [CrossRef]
9. Huang, J.; Chen, G.; Shu, L.; Wang, S.; Zhang, Y. An Experimental Study of Clogging Fault Diagnosis in Heat Exchangers Based on Vibration Signals. *IEEE Access* **2016**, *4*, 1800–1809. [CrossRef]
10. Li, G.; Chen, L.; Liu, J.; Fang, X. Comparative Study on Deep Transfer Learning Strategies for Cross-System and Cross-Operation-Condition Building Energy Systems Fault Diagnosis. *Energy* **2023**, *263*, 125943. [CrossRef]

11. Wang, K.; Chen, J.; Song, Z.; Wang, Y.; Yang, C. Deep Neural Network-Embedded Stochastic Nonlinear State-Space Models and Their Applications to Process Monitoring. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 7682–7694. [CrossRef] [PubMed]
12. Davoudi, E.; Vaferi, B. Applying Artificial Neural Networks for Systematic Estimation of Degree of Fouling in Heat Exchangers. *Chem. Eng. Res. Des.* **2018**, *130*, 138–153. [CrossRef]
13. Habbi, H.; Kidouche, M.; Kinnaert, M.; Zelmat, M. Fuzzy Model-Based Fault Detection and Diagnosis for a Pilot Heat Exchanger. *Int. J. Syst. Sci.* **2011**, *42*, 587–599. [CrossRef]
14. Liang, Y.; Zhu, L.; Wang, Y.; Wu, H.; Zhang, J.; Guan, J.; Wang, J. Fouling Prediction of a Heat Exchanger Based on Wavelet Neural Network Optimized by Improved Particle Swarm Optimization Algorithm. *Processes* **2024**, *12*, 2412. [CrossRef]
15. Tian, W.; Liu, N.; Sui, D.; Cui, Z.; Liu, Z.; Wang, J.; Zou, H.; Zhao, Y. Early Warning of Internal Leakage in Heat Exchanger Network Based on Dynamic Mechanism Model and Long Short-Term Memory Method. *Processes* **2021**, *9*, 378. [CrossRef]
16. Jallal, M.A.; Vallée, M.; Lamaison, N. Fouling Fault Detection and Diagnosis in District Heating Substations: Validation of a Hybrid CNN-Based PCA Model with Uncertainty Quantification on Virtual Replica Synthesis and Real Data. *Energy* **2024**, *312*, 133590. [CrossRef]
17. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 4768–4777.
18. Le, M.; Vu, H.H.T.; Wang, D.; Lee, J. Development of Electromagnetic Cylinder-Type Probe for Inspection of Heat Exchanger Tubes. *IEEE Trans. Magn.* **2022**, *58*, 4001609. [CrossRef]
19. Spiegel, C.; Kraut, M.; Rabsch, G.; Küsters, C.; Augustin, W.; Scholl, S. Fouling Detection in an Optically Accessible Microstructured Heat Exchanger. *Chem. Eng. Technol.* **2019**, *42*, 2067–2075. [CrossRef]
20. Bykov, S.P.; Inshakov, D.V.; Kuznetsov, K.A. Application of Acoustic Pulse Reflectometry in Nondestructive Testing of Heat Exchanger Pipes. *Chem. Pet. Eng.* **2015**, *50*, 670–674. [CrossRef]
21. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.