A Data Privacy Protection Diagnosis Framework for Multiple Machines Vibration Signals Based on a Swarm Learning Algorithm

Shilong Sun, *Member, IEEE*, Haodong Huang, Tengyi Peng, Changqing Shen, *Senior Member, IEEE*, and Dong Wang, *Member, IEEE*

***Abstract*—Bearing fault diagnosis is essential for monitoring rotating machinery and equipment operating conditions. They can also identify safety hazards and avoid economic losses promptly. However, there is a shortage of tagged fault data in most factories, and it is challenging to ask them to exchange labeled fault data, especially considering the data privacy needs of customers. This article proposes a swarm learning (SL) framework that combines adversarial domain networks with convolutional neural networks (CNNs) to address this problem. The framework regards every factory as an edge-computing node and solves labeled data insufficiency and privacy protection by fusing network parameters. First, a CNN is used to compute each node, and leaders are dynamically selected to merge model parameters during the training process. Second, an adversarial domain network minimizes the feature distribution variance between nodes. Finally, an SL algorithm was used to select virtual central nodes to determine the exchange process of the model parameters among different nodes. Four datasets were used to design the experiments and demonstrate the proposed approach's reliability. The experimental results show that the proposed framework can improve computational efficiency and reduce communication costs without relying on a central server. A final shared model can also achieve enhanced accuracy in fault diagnosis at each edge node.**

***Index Terms*—Adversarial domain network, data privacy protection, fault diagnosis, swarm learning (SL) algorithm, weight parameter fusion.**

Manuscript received 9 September 2022; revised 20 November 2022; accepted 19 December 2022. Date of publication 4 January 2023; date of current version 12 January 2023. This work was supported in part by the Natural Science Foundation of Guangdong Province under Grant 2021A1515110615, in part by the National Natural Science Foundation of China under Grant 52272440 and Grant 51975355, and in part by the Soochow Science and Technology Innovation Talents Program under Grant WC202168 and Grant ZXL2022027. The Associate Editor coordinating the review process was Dr. Arunava Naha. *(Corresponding author: Changqing Shen.)*

Shilong Sun, Haodong Huang, and Tengyi Peng are with the School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen 518055, China, and also with the Guangdong Key Laboratory of Intelligent Morphing Mechanisms and Adaptive Robotics, Shenzhen 518055, China (e-mail: sunshilong@hit.edu.cn).

Changqing Shen is with the School of Rail Transportation, Soochow University, Suzhou 215006, China, and also with Suzhou Boyata Industrial Internet Company Ltd., Suzhou 215100, China (e-mail: cqshen@suda.edu.cn).

Dong Wang is with the State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: dongwang4-c@sjtu.edu.cn).

Digital Object Identifier 10.1109/TIM.2023.3234035

I. INTRODUCTION

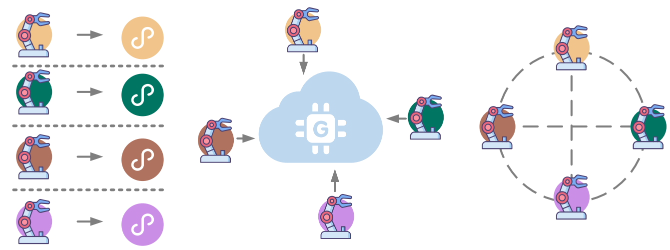
IN MODERN industries, data-driven methods [1], [2], [3] are significant for diagnosing mechanical component faults. They allow real-time monitoring of equipment operation, thus improving the reliability and safety of equipment operation and reducing the cost of maintenance. In recent decades, the rapid development of artificial intelligence (AI) algorithms has successfully applied fault diagnosis for crucial mechanical components, such as bearings, gearboxes, and motors. The rapidity and timeliness of intelligent diagnostic methods have led to their widespread use in the mechanical and electrical fields.

A data-driven model for diagnosing faults is a typical example of applying AI algorithms to mechanical fault diagnosis. For example, convolutional neural networks (CNNs), autoencoding, recurrent, and belief networks [1], [2], [3], [4]. Jia et al. [5] proposed a multi-scale residual attention CNN (MRA-CNN) fault diagnosis framework. Liu et al. [6] summarized AI algorithms to diagnose rotating machinery faults. The models are trained using input data, and the trained models are evaluated using test sets. Xu et al. [7] proposed a new depth transfer CNN (TCNN) framework for online fault diagnosis, which converts original signals into gray-scale images and uses these images to achieve online fault diagnosis. Peng et al. [8] proposed residual neural networks for label-free intelligent fault diagnosis of bearings, and Wang et al. [9] proposed an interpretable neural network for resonance localization to perform machine state detection. Von Hahn and Mechefske [10] used a disentangled-variational-autoencoder with a temporal CNN. This is an end-to-end deep learning (DL) method to trend and predict tool wear in the input and talent space. Yang et al. [11] proposed a transfer learning (TL) approach for bearing fault diagnosis, and this method utilizes laboratory diagnostic knowledge to diagnose the health status of an actual machine. Zhou et al. [12] proposed a multi-channel cutting force signal tool wear condition monitoring (TCM) detection method based on multi-scale edge-labeling graph neural network (MEGNN) and phase space reconstruction (PSR). This method outperforms several other CNN networks. These AI-based and data-based intelligent fault diagnosis models can achieve high accuracy due to a large training dataset and vast computing power. Therefore, the numerous training dataset and colossal computing power are prerequisites to realize diagnostic accuracy in some specific scenarios.

Although the AI-based diagnosis models have achieved high accuracies, some issues still exist, such as the low-quality data to train the AI model, inadequate data privacy protection, etc. High-quality data can be defined by those with sufficient labels, full-life long run time, noticeable fault features, and balanced data distribution. However, collecting a large amount of high-quality data under normal operating conditions is still a challenge for typical plants and factories. Hence, conducting intelligent fault diagnosis for these datasets presents some obstacles. The most parsimonious approach is expanding the dataset by using multiple similar machines to collect limited data, but gathering the data would require considerable time and labor. At the same time, as the value of industrial data increases and the demand for data privacy protection, the problem of "data silos" is forming in the real industry. Therefore, the lack of data volume further limits the development of DL-based intelligent diagnosis algorithms.

Therefore, establishing an approach and conducting a complete intelligent fault diagnosis without sharing local data are urgently needed. Federated learning (FL) can solve the data privacy protection issue. McMahan et al. [13] applied FL for the first time by leaving the training data distributed on local storage and learning shared model updates, which can be utilized for unbalanced and non-independent and identically distributed (IID) data distribution. FL allows local diagnostic models from different participants to communicate with each other, while their data is still stored locally, and data privacy is preserved. The participants participating in the joint training are called clients, and the trusted center that supports the clients for such communication is called the server. Zhang et al. [14] proposed a novel self-monitoring FL fault diagnosis method using fake data samples to learn the data structure thoroughly. Chen et al. [15] proposed a fault detection method based on a joint neural network. Because of its advantages, such as self-joint learning, this neural network has sufficient ability to transfer and adjust knowledge across domains. Lin et al. [16] proposed a CNN with a channel attention mechanism, divided the model layer into shallow and deep layers, and finally shared only shallow parameters. Chen et al. [17] designed a dynamic weighted average algorithm based on the maximum mean discrepancy (MMD) to integrate model parameters in FL. Wang et al. [18] proposed a novel federal TL (FTL) framework. And a joint minimum [federated minimax (FedMM)] algorithm for global model aggregation is introduced. Zhang and Li [19] proposed an FL method based on deep adversarial networks to keep similar data structures in distributed extracted features and solve the prediction consistency scheme. Zhang and Li [20] proposed an FL method based on the prior distribution to minimize the domain gap. The methods proposed above are all applications of FL in mechanical equipment fault diagnosis, which solves the problem of insufficient data and feature distribution in different fields. Nonetheless, the communication cost of FL is exceptionally high.

From the above survey, it can be concluded that FL can solve the problem of data privacy protection and insufficient data, but at the same time, it also has the problem of high communication costs. Swarm learning (SL) can solve the problems of insufficient data and high communication costs to some extent. Compared to FL, SL can reduce communication costs and protect data privacy. SL was proposed by Warnat-Herresthal et al. [21] and used for decentralized confidential clinic datasets; it has a similar role to FL but differs from it in structure. It does not need a central server to control and distribute all labels and features to nodes in the network. It enables every member of the network to share the information directly without needing an external central server [22]. Therefore, this structure is more stable and avoids the problem of excessive central server rights. Fig. 1 depicts the differences in network structure between local, federated, and SL algorithms.



(a) (b) (c)

Fig. 1. Schematic of the structure of the three methods. (a) Local learning: described in different places. (b) FL: data is held together by data contributors and a central parameter server coordinates parameter settings. (c) SL: no central server is needed.

To address the issues of insufficient data, data privacy protection, and reduced communication costs. This article proposes an SL approach based on CNN and domain adversarial neural networks (DANN). The CNN is used to train the local diagnosis model, the DANN is used to minimize the data distribution gap between clients, and the SL is utilized to select decision-makers, distribute model parameters, and be in charge of the model updates. This method is suitable for fault diagnosis of multiple mechanical parts in the case of insufficient data and fast diagnosis within privacy protection. The main innovations and contributions of this article are as follows.

1) A new local training method of fault diagnosis SL framework is constructed. The domain transfer problem is carefully considered during local training, where a DANN-based training strategy is used.
2) A new model parameter averaging algorithm is proposed to integrate models, unify model distribution, reduce communication costs, and converge faster than gradient transfer.
3) The experimental results from vibrational datasets show that the proposed method can handle signals from various devices and domains and protect their local data privacies. It has considerable potential in the application of other industrial scenarios.

The remainder of this article is organized as follows. Section II describes related work, including the CNN, DANN, and SL algorithms. Section III presents the procedure of the proposed method. Section IV indicates the design and validation of experiments, and Section V concludes this article.

II. Related Work

A. DANN

A feature distribution discrepancy generally exists among machines because their operating condition and fault types are different. This variance deteriorates the robustness of the diagnosis model. TL can utilize the knowledge learned in previous tasks for new tasks by extracting the domain invariance feature [23], which is widely used in fault diagnosis tasks. The essential part of TL is domain adaption, which deals with the inconsistent data distribution in the source and target domains. It aims to map the data in the source and target domains to the same feature subspace to allow the source domain feature discriminator to be applied to the target domain. MMD is one of the well-known domain adaptation methods widely used in bearing fault diagnosis. Lu et al. [24] added MMD terms to the loss function, achieving unlabeled target domain data classification. Yang et al. [11] placed MMD on each convolutional layer. They combined the proposed TL network with a pseudo-label training method, allowing to transfer of the diagnosis knowledge learned from laboratory bearings to locomotive bearings. The equation of MMD can be expressed as follows:

$$\mathrm{MMD}(P, Q) = \| E_{X\sim P}[\varphi(X)] - E_{Y\sim Q}[\varphi(Y)] \|_H \tag{1}$$

where P and Q are the distributions over sets X and Y. $\varphi(X)$ and $\varphi(Y)$ are the source and target data, respectively, the distribution distance between the source and target data in the original space is equivalent to the distance between the means of the projected source and target data in the embedded space. This distance is based on embedding probabilities in a reproducing kernel Hilbert space.

DANN is another popular domain adaptation method [25], [26]. Aside from the standard feed-forward feature extractor and label predictor, DANN contains a domain classifier that connects to the feature extractor via a gradient reversal layer. The domain classifier is utilized to distinguish the source of the input data. During backpropagation-based training, the gradient reversal layer multiplies the gradient by a certain negative constant. The training process must minimize the label prediction and domain classification losses. With the help of the domain classifier and gradient reversal layer, the feature distribution of all domains will be similar.

B. Swarm Learning

Swarm intelligence (SI) represents the collective behavior of decentralized, self-organizing systems. It is inspired by organisms, such as ant colonies, bird flocks, animal grazing, and bacterial growth. SI is currently used in various fields, such as library access, communication, and medical data classification. SL [27] is inspired by SI, which does not require a dedicated server like FL. The dedicated parameter servers are responsible for aggregating and distributing the local learning network model in FL. SL dispenses with the dedicated server, shares the local model parameter via the swarm network, and independently builds the local models from the private data of individual machines. Therefore, the user's data privacy is further protected, and the user's distrust of the dedicated server is dispelled. Moreover, the reticular structure of SL is more robust than the star structure of FL.

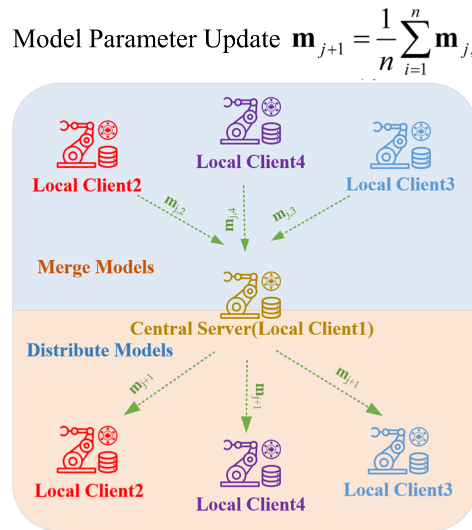


Fig. 2. Schematic of the SL algorithm. (Each epoch randomly selects a decision maker to collect, integrate, and distribute the models. The illustration takes the local client1 as the virtual central server as an example.)

Fig. 2 shows the schematic of the diagnosis structure based on the SL network. SL can authorize each participant (local client) to share and update the model, which means that each participant has the same ability as the central server of FL. Therefore, each participant can integrate information from multiple other edge computing node participants. In the training process, each participant should accomplish two steps. The first step is running the machine learning algorithm locally, responsible for iterating and updating the offline model. The second step is to update the information to the SL network that enables parameter interaction.

Furthermore, the swarm network authorizes a participant to collect model parameters from other participants to integrate the local model dynamically. For example, four clients participate in the SL network shown in Fig. 2. First, any client is selected randomly as the virtual central server to train the local model and distribute the model parameters to the three other clients. Second, the three other clients merge, update the model, and return their model parameters to the virtual central model. The first and second steps complete the whole communication process of model parameters. Compared with the data centralized training method, SL does not need to consider the privacy and security issues in the data sharing because only the model parameters are exchanged during the model training process. It also helps prevent data monopoly caused by data centralization. Accordingly, the SL method is highly suitable for the equipment intelligent diagnosis task of a factory cluster.

III. Procedure of the Proposed Method

This section specifies the proposed population learning network approach based on CNNs and domain adversarial.

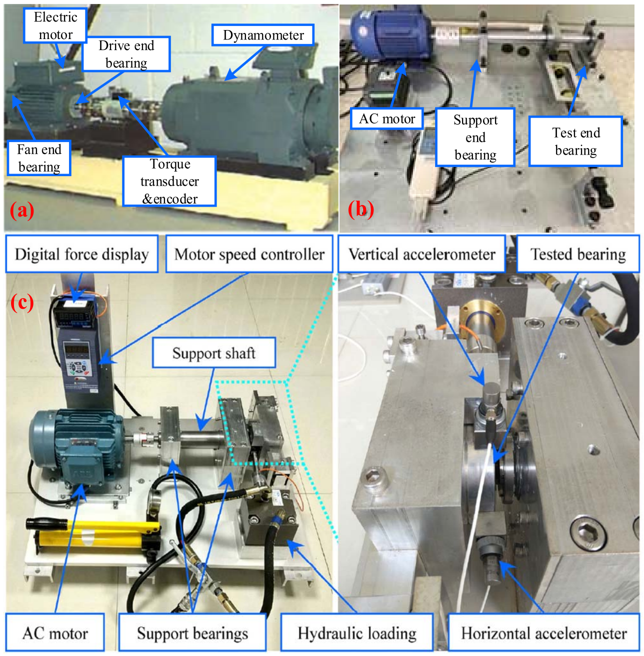


Fig. 3. Test rigs of the rotating machines in (a) CWRU, (b) HITsz, (c) XJTU, and (d) SCU.

TABLE I
DESCRIPTIONS OF FOUR DATASETS

Datasets	Motor load	Motor speed(rpm)	Fault diameter	Sampling frequency(Hz)
CWRU	1 HP	1772	7 miles	12k
HITsz	1 kN	1200	0.3mm	25.6k
XJTU-SY	10 kN	2400	\	25.6k
SCU	1 kN	896.1	0.3mm	10k

A. Datasets

In the process of SL, the locally stored data are different for various nodes, so each node has the same size dataset. The four types of bearing data, normal, inner ring failure, outer ring failure, and rolling body failure, also have the same amount of data in each node. The bearing ring stand is shown in Fig. 3. The working conditions and specific conditions of each data set are shown in Table I.

Dataset A (CRWU) [28]: The dataset for the first node is a publicly available dataset from Kate Western Reserve University.

Dataset B (HITsz): The dataset for the second node was obtained from our own laboratory collection.

Dataset C (XJTU-SY) [29]: The third node's dataset is from Xi'an Jiaotong University's public dataset.

Dataset D (SCU): This is from the Soochow University dataset.

B. Proposed Method

This section presents the SL algorithm based on CNNs and domain adversarial. Each node implements the machine learning algorithm locally. After one training session, the swarm network dynamically elects a decision maker who collects the other model parameters to integrate the model and then distributes it to the other nodes. A decision maker is selected after each training session until the end of the cycle to obtain the final model.

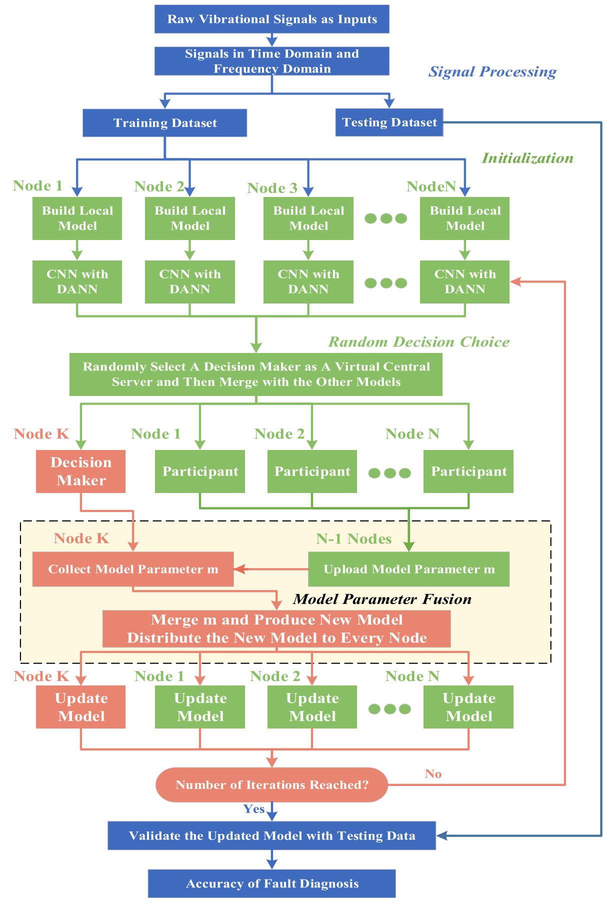


Fig. 4. Workflow of the proposed swarm intelligent fault diagnosis based on the SL network.

In the training process, the data at different nodes may be obtained under different working conditions and have various feature domain distributions. Consequently, the results are not reliable if the data are trained without domain adaption processing. Here, instead of reducing the gap between the source and target domains, we introduce domain adversarial to reduce the data "distance" between the domains where the four nodes are located. The workflow of the proposed swarm intelligent fault diagnosis based on the SL network is shown in Fig. 4.

C. Network Structure and Optimization

Due to DANN's use in the network, the loss function is no longer a simple cross-entropy loss function. It needs to consider the returned gradient due to the presence of a gradient inversion layer between the feature extractor and the domain discriminator in the network. First, for the optimization of the local model, the empirical classification error is assumed to be minimal, and the widely used cross-entropy loss function is used, for the local data $\{(x_i, y_i)\}_1^{n_s}$, as follows:

$$L_y = -\frac{1}{n_s}\sum_{i=1}^{n_s}\sum_{j=1}^{Nc} 1\{y_i = j\}\log\frac{e^{x_{i,j}^s}}{\sum_{k=1}^{Nc} e^{x_{i,j}^s}} \tag{2}$$

where n_s is the number of all labeled data, y_i is the label corresponding to the data, N_c value is the number of bearing categories, and $x_{i,j}^h$ is the jth neuron of the fully connected layer in the labeled classifier. L_y is the loss function of the node in training the CNN network locally.

Owing to the gradient inversion layer, the loss function should be added to the part of the inversion back to the gradient

$$L_d = -\frac{1}{n_d}\sum_{i=1}^{n_d}\sum_{j=1}^{N_d} 1\{d_i = j\}\log\frac{e^{x_{i,j}^d}}{\sum_{k=1}^{N_d} e^{x_{i,k}^d}} \tag{3}$$

where d_i refers to the label of the domain, $x_{i,j}^d$ refers to the jth neuron of the fully connected layer in the domain discriminator, and N_d is the number of the domains.

Summarizing (2) and (3) yields the final loss L_{obj} as follows:

$$\min L_{\text{obj}} = L_y + \lambda L_d \tag{4}$$

where λ is a hyper-parameter between the two objectives to tradeoff source risk and DANN. Therefore, it is an artificially set regularization parameter and λL_d is to prevent overfitting. The range λ is taken as (0, 1). In this article, λ is taken as 0.1.

D. Merge Model

In integrating the model, a decision maker is dynamically selected to collect the parameters of other nodes. This article proposed two approaches to conduct the model fusion process, one is called model parameter mean [weight mean (WM)], and the other one is the weight (loss function) multiplied by model parameters [weight dynamic fusion (WDF)]. Since the local training model is a CNN structure, then the WM and WDF are worked both at the convolutional layer and the fully connected layer. In a word, the model parameter updating process only occurs at the convolutional and fully connected layers. Fig. 5 displays the fusion process of two model parameter updating approaches.

The WM is inspired by the federated average algorithm. The difference is that federal learning transmits the gradient while we only deliver the weight and bias here. The formula for WM is as follows:

$$\mathbf{m}_{j+1} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{m}_{j,i} \tag{5}$$

where $\mathbf{m}_{j+1}$ refers to the integrated model parameters, i refers to the number of nodes, and $\mathbf{m}_{j,i}$ refers to the model parameters of the node obtained after the jth round of training. $\mathbf{m}$ contains two components, weight $\mathbf{w}$, and bias $\mathbf{b}$.

The model integration process draws on the experience of the federal averaging algorithm where $\mathbf{w}$ and $\mathbf{b}$ are pooled for all clients, and the parameters are averaged and calculated by the decision maker and distributed to each client. This model-averaging approach reduces communication costs and

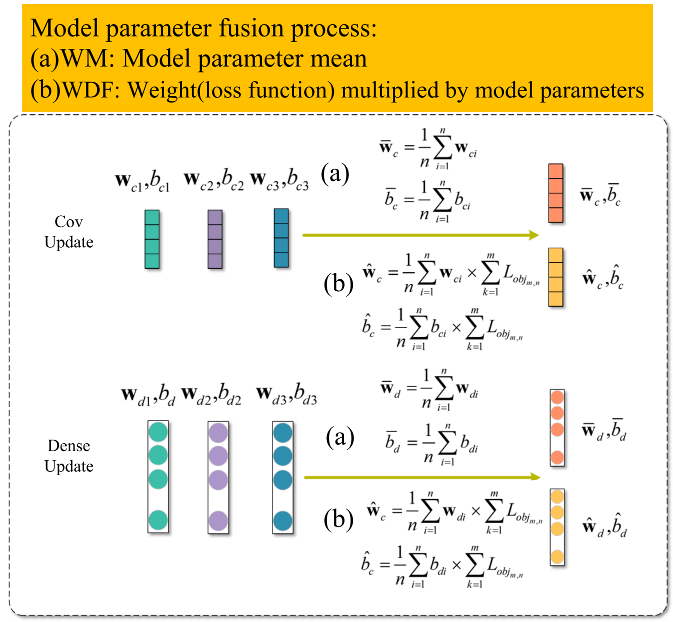


Fig. 5. Schematic of the fusion process of two model parameter updating approaches. This figure includes the parameter update method of the convolutional layer and the fully connected layer.

converges faster than the gradient transfer. The complete pseudo-code for the merging model process is given in Algorithm 1.

Algorithm 1 Swarm Averaging. The n Clients Are Indexed by i; E Is the Number of Local Epochs

I. Upload model parameters

Initialize w_1, b_1

For each node, $i = 1, 2, \ldots, n$ do

Select $i = k_{th} \subset (1, 2, \ldots, n)$ as the decision maker

$m_{j,k_{th}} = (w_{k_{th}}|b_{k_{th}})$ a random node k_{th} of $n$clients

Train the other $n-1$ models

$$m_{j,n-1} = (w_1, w_2, \ldots, w_{n-1}|b_1, b_2, \ldots, b_{n-1})$$

II. Virtual server update

All model parameters averaging

$$\mathbf{m}_j = \frac{1}{n}\left(\sum_{i=1}^{n-1} m_{j,i} + m_{j,k_{th}}\right)$$

III. Distribute updated parameter

For each epoch, $j = 1, 2, \ldots, E$ do

$$\mathbf{m}_{j+1} = \mathbf{m}_j$$

Return $\mathbf{m}_{j+1}$ to all nodes

The WDF aims to fuse model weights by multiplying the loss function dynamically

$$\mathbf{m}_{j+1} = \sum_{i=1}^{m}\left(\mathbf{m}_{j,i} \times \sum_{k=1}^{n} L_{\text{obj}_{j,i,k}}\right) \tag{6}$$

where $\mathbf{m}_{j+1}$ refers to the integrated model parameters, $\mathbf{m}_{j,i}$ refers to the model parameters of the node obtained after the jth round of training, and $L_{\mathrm{obj}_{i,j,k}}$ refers to the loss function of the ith node obtained at the kth batch of the jth round of training.

The transmitted parameters refer to the parameters trained by the whole model, including the weights and biases of all convolutional and fully connected layers. The central node adds the model parameters of each node and divides them by the number of nodes. Given that the model of each node is the same, only the trained parameters are different. Therefore, the summed and averaged parameters of each node at the same position are the new parameters of the corresponding position of the new model. For example, there are three convolutional layers in the model, and there are 64 filters in the first layer. The filter length is 32, so there are 32×64 weights and 64 biases. We add 2112 parameters of each node and divide them by the number of nodes, which are the parameters of the first convolutional layer of the new model.

IV. Experimental Results and Discussion

The assumption is made that we have four clients. The element and node on the SL network are clarified clearly. Then, the experiments are designed and conducted with four nodes.

A. Data Preprocessing

In the experimental process of SL, there are four nodes. Given that the data of SL are stored in local storage for training, each node's data cannot be overlapped. The training dataset for node 1 is dataset A, node 2 is dataset B, for node 3 is dataset C, and for node 4 is dataset D. There are four types of bearing health status: normal, inner race failure, outer race failure, and rolling element failure. The dataset for each node has 404 sets of data, and the length of each set of data is 1200 points. For each node, 200 are used as training datasets and 204 are used as test datasets. For each fault type, 50 sets of data are used to train the diagnosis model, while 51 sets of data are used to test this model. The network structure of CNN is shown in Table II.

B. Visualization of Experimental Results

In this experiment, the model is trained for each node, and the same model is finally integrated using CNN for feature extraction and domain partitioning through domain adversarial. The results are visualized through principal component analysis (PCA) dimensionality reduction. The hyperparameters of the experiments are listed in Table III.

The effectiveness of CNN and the DANN-based SL network approach is demonstrated through visualization. PCA and t-distributed stochastic neighbor embedding (t-SNE) reduce the dimensionality of the learned high-dimensional features to a 2-D plane to cluster the conversational data. The downscaling data plots for the four nodes are shown in Fig. 6. The PCA and t-SNE dimensionality reduction methods clearly show that each fault type is clustered together. The experimental results demonstrate that the proposed method can achieve a cross-domain and cross-device diagnosis.

TABLE II
Parameter Setting of the CNN

Layer	Tied parameter	Activation function
Input	/	/
Convolutional layer 1	Kernel size: 64 Filters：32 Stride:4	ReLU
Pooling layer 1	Pool_size: 5 Stride: 5	/
Convolutional layer 2	Kernel size: 8 Filters：64 Stride:4	ReLU
Pooling layer 2	Pool_size: 3 Stride: 3	/
Convolutional layer 3	Kernel size: 2 Filters:128 Stride:1	ReLU
Pooling layer 3	Pool_size: 3 Stride: 3	/
Flatten	/	/
Dense 1	Units:180	ReLU
Dense 2	Units:4	SoftMax

TABLE III
Parameters Used in This Experiments

parameter	value	parameter	value
Batch size	24	learning rate	0.0002
epoch	180	λ	0.01
Ntrain	50	Ntest	51

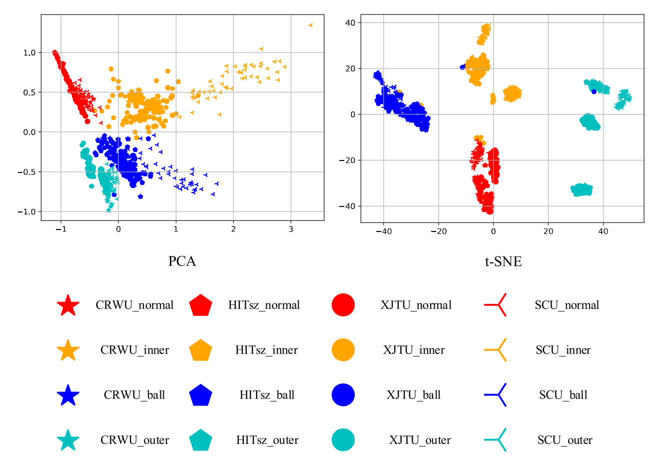


Fig. 6. Feature dimensionality reduction plots for four datasets under the SL framework. The results of the model under different dimensionality reduction methods are given.

C. Experimental Design

We verify the superiority of the proposed approach through the following four experimental designs. Experiment I is to study the influence of the size of the training dataset on the experimental results. Experiment II is a comparison of the three methods. Experiment III compares SL and local learning.

Experiment I: This experiment investigates the effect of training set size on joint fault diagnosis under the SL

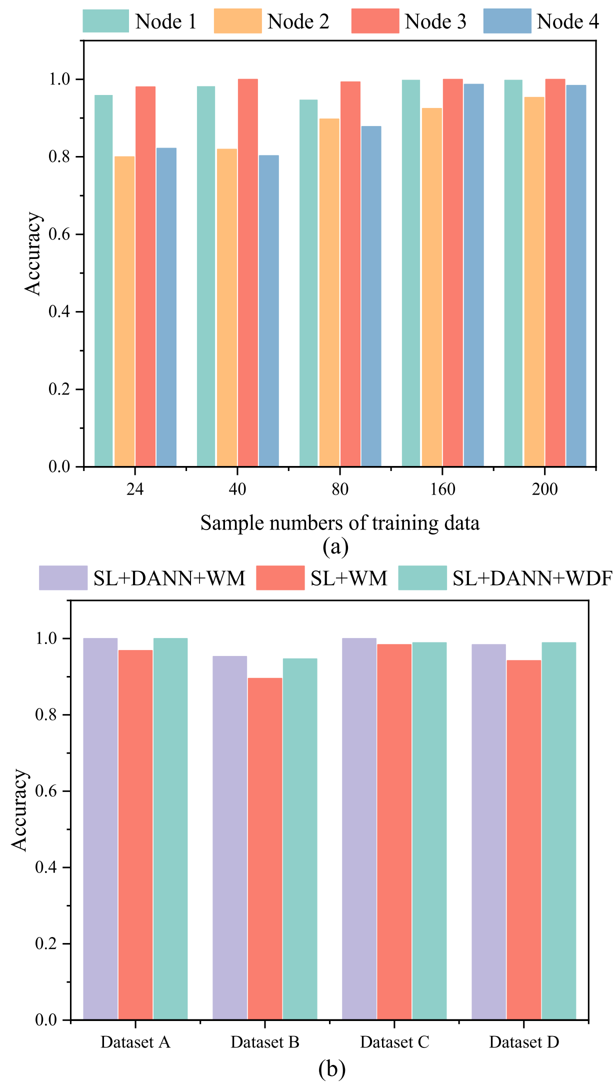


Fig. 7. (a) EXP I: Accuracy of the four nodes under different sample numbers of the training data. (b) EXP II: SL + DANN + WM; SL + WM; and SL + DANN + WDF.

framework. The time domain signal of the vibration signal is used in this experiment. Considering that the batch size in the training process is 24, we set the training data size to 24, 40, 80, 160, and 200. The experimental results are shown in Fig. 7(a). With the increase in training data, the accuracy of the final test also increases. When the training dataset is only 24 (there are six data in each health state), each node can achieve more than 80% accuracy. This result shows that the proposed network can achieve a certain diagnosis accuracy when the number of samples is small.

Experiment II: This experiment verifies the superiority of the proposed method. The temporal vibration signal is used in this experiment. Experiment A (SL + DANN + WM) is the proposed method in this article; Experiment B (SL + WM) is Experiment A without using DANN; Experiment C (SL + DANN + WDF) is a fusion method that adopts a dynamic fusion of model parameters. This fusion method regards the loss function obtained by each epoch as the new weight, multiplying them with the modal parameters

TABLE IV
SL VERSUS LOCAL LEARNING FOR COMPARISON (FOUR NODES)

Accuracy score (%)		Swarm	Model 1	Model 2	Model 3	Model 4
Time Domain	Node 1	100	100	26.56	26.56	26.56
	Node 2	95.31	17.71	95.31	26.56	26.56
	Node 3	100	1.04	38.54	100	27.08
	Node 4	98.44	0	10.94	44.79	100
Frequency Domain	Node 1	100	100	14.58	0	35.42
	Node 2	100	26.56	100	26.56	36.46
	Node 3	100	0.52	30.73	100	45.31
	Node 4	100	6.77	22.40	52.60	100

and then adding them together as the new model parameter. The weights change continuously as the training process progresses. Fig. 7(b) shows three different methods for the diagnosis accuracy: 1) SL + WM; 2) SL + DANN + WM; and 3) SL + DANN + WDF. The experimental results show that DANN can improve the problem of data in different domains, thereby improving the performance of the whole SL network framework. Comparing the two methods of model parameter fusion, the WM method achieves better performance than WDF. WDF is equivalent to adding weight to the model parameters of each node, which not only increases the amount of calculation but also reduces the diagnostic accuracy.

Experiment III: This experiment compares local learning with SL and establishes two groups of experiments. In this experiment, both time domain and frequency domain signals are studied. The first group of experiments includes the method proposed in this article. The second group of experiments is to remove the process of model integration. Each node only uses its data for local learning and shares the model it trains with other nodes. Finally, the accuracy of the two methods is compared. The training data for this part is still 200 groups of data.

Table IV shows the result of SL and local learning in the time domain and frequency domain (Experiment III). It can be seen that the frequency domain is equivalent to performing feature extraction on the original data once, and obtains a higher effect. But if the network parameters are not fused and only local training is performed, the trained network results only apply to the nodes themselves. When the model of one node is tested on another node, the effect is less good. Therefore, the diagnostic accuracy of the co-training framework will not be very high. Overall, the results obtained by the proposed method in this experimental design are similar to those obtained by training our models locally. It shows that the proposed method can solve the cross-domain problem and achieve an accuracy that does not lose the local training.

Experiment IV: In the above three groups of experiments, the datasets trained by the four nodes are equal, and the size of each node's four health state datasets is the same. Now, we consider having different dataset sizes for each node for the four types of data. This experiment loses a portion of one kind of data from each node, and each node loses a different type of data. The experiment setup follows: node 1 is missing

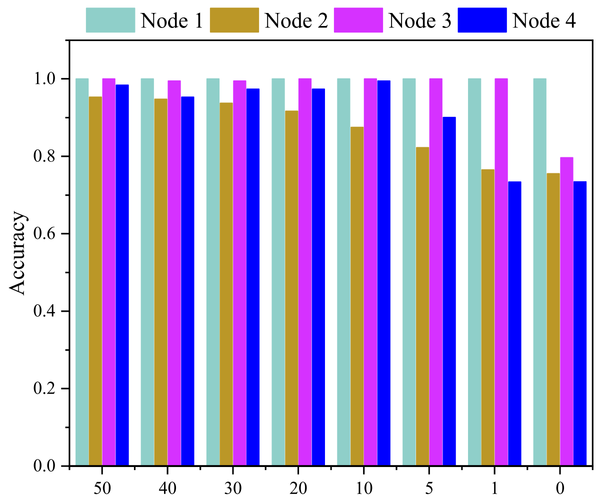


Fig. 8. Experimental results of missing training data. Different colors in the figure represent various nodes, and the missing data types of each node differ.

TABLE V
DATASET PROPORTION SETTING FOR EACH EXPERIMENTAL GROUP IN EXPERIMENT IV

Group type		50	40	30	20	10	5	1	0
Node 1	N	50	50	50	50	50	50	50	50
	IR	50	40	30	20	10	5	1	0
	R	50	50	50	50	50	50	50	50
	OR	50	50	50	50	50	50	50	50
Node 2	N	50	50	50	50	50	50	50	50
	IR	50	50	50	50	50	50	50	50
	R	50	40	30	20	10	5	1	0
	OR	50	50	50	50	50	50	50	50
Node 3	N	50	50	50	50	50	50	50	50
	IR	50	50	50	50	50	50	50	50
	R	50	50	50	50	50	50	50	50
	OR	50	40	30	20	10	5	1	0
Node 4	N	50	40	30	20	10	5	1	0
	IR	50	50	50	50	50	50	50	50
	R	50	50	50	50	50	50	50	50
	OR	50	50	50	50	50	50	50	50

bearing inner ring failure data, node 2 is missing rolling body data, node 3 is missing outer ring failure data, and node 4 is missing normal data. Each node has the bearing data in normal conditions; however, to verify the method proposed in this article, node 4 will be missing some normal data. Additionally, each node has a data size of 50 for each type of bearing health status, which is now adjusted to 40, 30, 20, 10, 5, 1, and 0. The experimental results of imbalanced training data are shown in Fig. 8. The detailed data division for Experiment IV is shown in Table V. *N* refers to normal, IR refers to bearing inner race fault, *R* refers to rolling element fault, and OR refers to outer race fault.

In this experiment, the failure data of each node is unbalanced. The experimental results show that when the missing reaches 10, the fault diagnosis accuracy of each node is still above 90%. At that time, when the number of missing data for faults reached 10, the fault accuracy of each node was above 80%. When the amount of missing data increases further, the overall accuracy does not perform well. This shows that the network can be applied to a situation where there is not much missing fault data.

V. CONCLUSION

This article proposes a data privacy protection diagnosis framework based on an SL algorithm for multiple machine vibration signals. This diagnosis framework not only allows data privacy protection but also reduces communication costs and realizes a fast, accurate diagnosis. The summary can be listed as: 1) multiple machines fault diagnosis model; 2) a novel model parameter updating approach for reducing the communication cost; 3) a data privacy protection mechanism for diagnosis framework; and 4) solving the problem of data cross-domain and cross-device. Therefore, the proposed method offers a promising fault diagnosis approach for multiple mechanical components diagnosis, which fully utilizes the data in each node to learn the features while ensuring the privacy of the data.

It should be pointed out that despite the satisfactory results, there is still a gap between the proposed method and expectations in the problem of extremely missing data. Further research work will be carried out to develop effective TL techniques to address the challenging extreme data imbalance problem in SL.

REFERENCES

[1] F. Jia, Y. Lei, N. Lu, and S. Xing, "Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization," *Mech. Syst. Signal Process.*, vol. 110, pp. 349–367, Sep. 2018.

[2] S. Haidong, J. Hongkai, L. Xingqiu, and W. Shuaipeng, "Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine," *Knowl.-Based Syst.*, vol. 140, pp. 1–14, Jan. 2018.

[3] H. Liu, J. Zhou, Y. Zheng, W. Jiang, and Y. Zhang, "Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders," *ISA Trans.*, vol. 77, pp. 167–178, Jun. 2018.

[4] H. D. Shao, H. Jiang, H. Zhang, and T. Liang, "Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2727–2736, Aug. 2018.

[5] L. Jia, T. W. S. Chow, Y. Wang, and Y. Yuan, "Multiscale residual attention convolutional neural network for bearing fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 3519413.

[6] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mech. Syst. Signal Process.*, vol. 108, pp. 33–47, Aug. 2018.

[7] G. Xu, M. Liu, Z. Jiang, W. Shen, and C. Huang, "Online fault diagnosis method based on transfer convolutional neural networks," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 2, pp. 509–520, Feb. 2020.

[8] T. Peng, C. Shen, S. Sun, and D. Wang, "Fault feature extractor based on bootstrap your own latent and data augmentation algorithm for unlabeled vibration signals," *IEEE Trans. Ind. Electron.*, vol. 69, no. 9, pp. 9547–9555, Sep. 2022.

[9] D. Wang, Y. Chen, C. Shen, J. Zhong, Z. Peng, and C. Li, "Fully interpretable neural network for locating resonance frequency bands for machine condition monitoring," *Mech. Syst. Signal Process.*, vol. 168, Apr. 2022, Art. no. 108673.

[10] T. Von Hahn and C. K. Mechefske, "Self-supervised learning for tool wear monitoring with a disentangled-variational-autoencoder," *Int. J. Hydromechatron*, vol. 4, no. 1, pp. 69–98, 2021.

[11] B. Yang, Y. Lei, F. Jia, and S. Xing, "An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings," *Mech. Syst. Signal Process.*, vol. 122, pp. 692–706, May 2019.

[12] Y. Zhou et al., "A new tool wear condition monitoring method based on deep learning under small samples," *Measurement*, vol. 189, Feb. 2022, Art. no. 110622.

[13] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-efficient learning of deep networks from decentralized data," 2016, *arXiv:1602.05629*.

[14] W. Zhang, X. Li, H. Ma, Z. Luo, and X. Li, "Federated learning for machinery fault diagnosis with dynamic validation and self-supervision," *Knowl.-Based Syst.*, vol. 213, Feb. 2021, Art. no. 106679.

[15] H. Chen, Z. Chai, B. Jiang, and B. Huang, "Data-driven fault detection for dynamic systems with performance degradation: A unified transfer learning framework," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 3504712.

[16] J. Lin, J. Ma, and J. Zhu, "Hierarchical federated learning for power transformer fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 3520611.

[17] J. Chen, J. Li, R. Huang, K. Yue, Z. Chen, and W. Li, "Federated transfer learning for bearing fault diagnosis with discrepancy-based weighted federated averaging," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 3514911.

[18] Y. Wang, J. Yan, Z. Yang, Y. Dai, J. Wang, and Y. Geng, "A novel federated transfer learning framework for intelligent diagnosis of insulation defects in gas-insulated switchgear," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 3517711.

[19] W. Zhang and X. Li, "Federated transfer learning for intelligent fault diagnostics using deep adversarial networks with data privacy," *IEEE/ASME Trans. Mechatronics*, vol. 27, no. 1, pp. 430–439, Feb. 2022.

[20] W. Zhang and X. Li, "Data privacy preserving federated transfer learning in machinery fault diagnostics using prior distributions," *Struct. Health Monitor.*, vol. 21, no. 4, pp. 1329–1344, 2021.

[21] S. Warnat-Herresthal et al., "Swarm learning for decentralized and confidential clinical machine learning," *Nature*, vol. 594, no. 7862, pp. 265–270, 2021.

[22] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE ICNN*, vol. 4, Nov./Dec. 1995, pp. 1942–1948.

[23] M. Sun, H. Wang, P. Liu, S. Huang, P. Wang, and J. Meng, "Stack autoencoder transfer learning algorithm for bearing fault diagnosis based on class separation and domain fusion," *IEEE Trans. Ind. Electron.*, vol. 69, no. 3, pp. 3047–3058, Mar. 2022.

[24] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, and T. Zhang, "Deep model based domain adaptation for fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2296–2305, Mar. 2016.

[25] Y. Ganin et al., "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2030–2096, 2016.

[26] K. Zhao, H. Jiang, K. Wang, and Z. Pei, "Joint distribution adaptation network with adversarial learning for rolling bearing fault diagnosis," *Knowl.-Based Syst.*, vol. 222, Jun. 2021, Art. no. 106974.

[27] S. Warnat-Herresthal et al., "Swarm learning for decentralized and confidential clinical machine learning," *Nature*, vol. 594, no. 7862, pp. 265–270, 2021.

[28] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," *Mech. Syst. Signal Process.*, vols. 64–65, pp. 100–131, Dec. 2015.

[29] B. Wang, Y. Lei, N. Li, and N. Li, "A hybrid prognostics approach for estimating remaining useful life of rolling element bearings," *IEEE Trans. Rel.*, vol. 69, no. 1, pp. 401–412, Mar. 2020.

Shilong Sun (Member, IEEE) received the Ph.D. degree from the City University of Hong Kong, Hong Kong, in 2018.

He is currently an Assistant Professor with the Harbin Institute of Technology, Shenzhen, China. His current research interests include vibration energy harvesting design, fault diagnosis and prognosis, decision-making with artificial intelligence (AI), and deep learning (DL) for industrial data. Now, he focuses on the remaining equipment life estimation research with DL and smart energy harvesting techniques.

Haodong Huang was born in Jingmen, China. He received the B.S. degree in mechanical engineering from the School of Mechanical Engineering and Automation, University of Northeast of China, Shenyang, China, in 2021. He is currently pursuing the M.S. degree in mechanical engineering with the School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen, China.

His current research interests include mechanical fault diagnosis and condition monitoring.

Tengyi Peng was born in Shaoguan, China. He received the B.S. degree in mechanical design manufacture and automation from the School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, China, in 2020. He is currently pursuing the M.S. degree in mechanical engineering with the School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen.

His current research interests include mechanical fault diagnosis and condition monitoring.

Changqing Shen (Senior Member, IEEE) received the B.S. and Ph.D. degrees in instrument science and technology from the University of Science and Technology of China, Hefei, China, in 2009 and 2014, respectively, and the Ph.D. degree in systems engineering and engineering management from the City University of Hong Kong, Hong Kong, in 2014.

He is currently a Professor with the School of Rail Transportation, Soochow University, Suzhou, China, and also the Founder of Suzhou Boyata Industrial Internet Company Ltd., Suzhou. His current research interests include signal processing and machine learning-based fault diagnosis.

Dr. Shen is an Associate Editor of the IEEE OPEN JOURNAL OF INSTRUMENTATION AND MEASUREMENT.

Dong Wang (Member, IEEE) received the Ph.D. degree from the City University of Hong Kong, Hong Kong, in 2015.

He was a Senior Research Assistant, a Post-Doctoral Fellow, and a Research Fellow with the City University of Hong Kong. He is currently an Associate Professor with the Department of Industrial Engineering and Management, Shanghai Jiao Tong University, Shanghai, China, where he is also with the State Key Laboratory of Mechanical System and Vibration. His current research interests include sparse and complex measures, signal processing, prognostics and health management, condition monitoring and fault diagnosis, statistical learning and machine learning, statistical process control, and nondestructive testing.

Dr. Wang is an Editorial Board Member of *Mechanical Systems and Signal Processing.* He is an Associate Editor of the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, *Measurement*, IEEE SENSORS JOURNAL, and *Journal of Dynamics Monitoring and Diagnostics*.