

Original Article

Federated Learning-Based Futuristic Fault Diagnosis and Standardization in Rotating Machinery

Vijayalakshmi K¹, Amuthakkannan Rajakannu², Ramachandran KP³, Sri Rajkavin A V⁴

¹Department of Electrical and Communication Engineering, National University of Science and Technology, Muscat 130, Oman.

²Department of Mechanical and Industrial Engineering, National University of Science and Technology, Muscat 130, Oman.

³Deanship of Graduate Studies and Research, National University of Science and Technology, Muscat 130, Oman.

⁴College of Design and Engineering, Department of Electrical and Computer Engineering, National University of Singapore, Singapore.

¹Corresponding Author : vijayalakshmi@nu.edu.om

Received: 10 July 2024

Revised: 20 August 2024

Accepted: 11 September 2024

Published: 30 September 2024

Abstract - In the industrial sector, intelligent sensors in fault diagnosis are becoming more critical in recent technological improvements. The prediction accuracy can be enhanced in fault diagnosis using 3-dimensional, sequential, real-time, and image data. Sensors that capture the vibration, sound, and image data are more critical in predicting unbalancing, tool wear, crack, misalignment, etc, in the rotating machinery to increase productivity and to provide an effective maintenance management system. Due to the fast development of industry 4.0 techniques, monitoring of mechanical machinery is experiencing explosive growth and getting more attention in the area of Fault Diagnosis (FD). Machine learning and Deep learning methods give promising results and accuracy in predicting faults in rotating machinery on shop floors. The success of AI-based models is due to the availability of comprehensive labeled data. Federated Learning (FL) is the machine learning subfield aiming to train an algorithm with a heterogeneous dataset. Data transmission from local facilities to a central server in AI models creates data privacy and security issues. Heterogeneous data analysis is a complicated process in predicting the machine fault in the central server because of millions of data during real-time condition monitoring. Decentralized data handling and analysis is mandatory in condition monitoring because of heterogeneity in data processing, data privacy, and security advantages. The application of federated learning in fault diagnosis has been getting more attention in recent days, and this study is a review of FL applications that address fault diagnosis in rotating machinery in the first phase. A comparison between the types of FL approaches in FD, and the use of aggregation algorithms and their applications will also be discussed in Phase 1. In phase 2, a novel methodology has been proposed using Federated learning to diagnose rotating machinery faults. The proposed method, FLOACOS, addresses how prediction challenges are solved using federated learning approaches by optimizing and standardizing the data at local facilities. This work will be helpful for future condition monitoring researchers and gives an overview of a novel method of the FL technique used in predicting faults and the progress made in the maintenance management of rotating machinery.

Keywords - Fault diagnosis, Federated Learning, Internet of Things, Machine learning, Rotating machinery, Condition monitoring.

1. Introduction

The application of AI models in fault diagnosis shows a promising result in maintenance management with more accuracy when using the data set collected across several industrial environments. The data collected from the sensor is processed in a centralized server, and a large amount of data has been stored in the cloud. Even though data sharing from sites to servers has many benefits, data privacy and security are big questions. To manage data privacy issues, federated learning will provide a solution for data collected from rotating machines, allowing for distributed training on edge

devices closer to the sensors or sites [1]. Vibration data for rotating machinery, which are predicted from petroleum Industries, Aircraft industries, Automobile industries, food factories, etc., are more confidential, and transferring through unreliable network connections to the server machines may give data privacy and security issues. The present fault diagnosis techniques will not address data security, economic conflict, the laws related to network issues, etc. So, we need to develop an intelligent diagnosis model with better network and server security performance. Weihua Li et al. [2] proposed a Clustering Federated Learning (CFL) method for bearing



fault diagnosis to address these issues. In the first phase of this work, a deep neural network with a self-attention mechanism is developed using a convolutional pipe to capture local and global information from the sensor output. Then, the clustering federated learning is used to collect data from different equipment unsupervised. The CFL model is fully utilized to use this data for fault diagnosis. Therefore, it is inferred that federated learning will improve privacy maintenance and diagnostic accuracy in rotating machinery. Rotating machinery in manufacturing, turbines, compressors, motors, etc., produces vast amounts of sensor data, typically distributed across different locations that may belong to other entities. FL allows the training of these entities without having a centralized computer or data analysis. In this type, each location trains a local model using its data and then distributes it to a central server. Typically, various sensors track parameters like vibration, temperature, pressure, and acoustic emissions in rotating machinery. The goal is to detect early signs of imbalance, wear, faults, misalignment, or anomalies before they lead to significant failures.

The standard method for fault diagnosis using FL involves a unified learning model as initial work. In the following work, the source domain of the client place will train the model using their data. After decentralized training, each client will upload their trained model to the server and then send it to the target client for further diagnosis. This method assumes that all client devices are working well and that target and source domain data are independently distributed. The disadvantage of this approach is that equipment variation will present if the independent node differs in working conditions. The data may not be independent and identically distributed, making FL inapplicable. An optimal aggression algorithm can be used in federated learning to overcome this.

Since there are many different conditions in the rotating machinery, this may give heterogeneous fault data, reducing the efficiency of federated learning. Researchers have proposed several methods to improve the efficiency of FL. Models like local model optimization, dynamic adaptive weight adjustments, weighted aggregation based on similarity, connecting two heterogeneous data using prior distribution, deep adversarial learning, asynchronous decentralized FL, self-attention mechanism, federated learning methods based on a variational auto-encoder, etc, are some of the methods to improve the efficiency of FL.

The novel approach in the proposed FL contains FLOACOS, which means managing the Origination of fault data, acquiring fault data, classifying fault data, optimizing fault data, and standardizing fault data in the local site itself before communicating to the central computer. 4IR techniques, including AI and IoT, have been used wherever needed in these phases.

2. Literature-Related AI-Based Machine Fault Diagnosis Models

Yaguo Lei et al. (2020) [3] reviewed the application of machine learning techniques in fault diagnosis and further explained the profound learning advantages in condition monitoring of machines in industries. They have discussed the challenges and roadmap for intelligent fault diagnosis and transfer learning on the topics of why to transfer”, “what is a transfer,” and “how to transfer.” Liu et al. [4] reviewed the applications of artificial intelligence in fault diagnosis of rotating machines and mainly focused on applications of traditional machine learning. Yaguo Lei and Ming J. Zuo (2009) applied a Weighted K Nearest Neighbor (WKNN) classification algorithm to identify the gear crack levels. This work captured vibration signals due to the gear crack under different loads and motor speeds. The K nearest neighbor method was applied to identify the gear crack levels, and the obtained results demonstrate its effectiveness [5]. B. Samantha predicted faults in gears using artificial neural networks and Support Vector Machines (SMVs). Time domain signals of rotating machines with non-defective and defective gears are used for feature extraction. The extracted features are used as input for ANN and SVMs [6].

A study is presented to compare the performance of gear fault detection using Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs). The time-domain vibration signals of a rotating machine with standard and defective gears are processed for feature extraction. The extracted features from original and preprocessed signals are inputs to both classifiers based on ANNs and SVMs for two-class (normal or fault) recognition. For most cases, SVMSVM's classification accuracy was better than ANN [7]. So, using modern technologies, fault diagnosis researchers investigate many methods, including machine learning, deep learning, digital twins, federated learning, etc. Since the traditional methods are less accurate, researchers are moving towards industry 4.0 technologies. Rui Zhao et al. [8] have surveyed in their paper and concluded, “It is believed that deep learning will have a more and more prospective future impacting machine health monitoring, especially in the age of big machinery data.” So, it indicates that deep learning is the future for precise monitoring machine and tool conditions. Lang Dai developed an Improved Deep Learning Model for Online Tool Condition Monitoring Using Output Power Signals. The sensor output mounted on the cutting tool holder during operation is used for further analysis. This data is analyzed using Long Short-Term Memory (LSTM) under a deep learning algorithm. Qun wang et al. [9] have done an overview of Tool Wear Monitoring Methods Based on Convolutional Neural Networks. The authors concluded that applying convolution neural networks in tool wear and condition monitoring is feasible and reliable. They have added that the convolution neural network can improve the prediction accuracy, which is an excellent significance of the

CNN technology. Convolutional Neural Networks (CNN), coming under deep learning, is a typical data-driven fault diagnosis method that extracts features from signals, images, or datasets using convolutional layers, then pooling and fully connected layers for tasks like image classification. In condition monitoring, researchers have recently applied CNNs to fault diagnoses of rotating machinery. Yao et al. [10] used an acoustic signal predicted using an acoustic emission sensor. They used CNN based on a multiscale dialog learning structure and attention mechanisms for gear fault prediction. Zhang, W et al. applied DCNN (Deep Convolutional Neural Networks) for bearing fault diagnosis under different operating loads [11]. Abdelmaksoud M et al. proposed a (CNN) model to diagnose induction motor faults at the motmotor's starting time. The model detects faults, such as locked rotor, overload, voltage imbalance, overvoltage, and Undervoltage under heavy, medium, and light load conditions [12]. The above methods concentrated the AI models in the rotating machinery fault detection, in which no model has been discussed regarding data privacy and data security, which have been highly demanding in recent periods due to cyber-attacks, data breaches, or data theft.

3. Review in Background Theory and the Challenges of Federated Learning

Federated learning is a machine learning technique that allows multiple institutions or devices to train a model while keeping data decentralized collaboratively. This approach is particularly relevant in fault diagnosis of rotating machinery due to its ability to handle data privacy and distributed data sources effectively. Lu, H et al. (2024) discussed Federated learning with uncertainty-based client clustering for fleet-wide fault diagnosis using their novel idea of federated learning, which is called FedSNGP. In this paper, the result of FedSNGP was compared with Fedavg and Fedcos, and the advantages and limitations of the FL were discussed [13]. Zhou F et al. (2024) paper aims to design an imbalanced FL framework based on dynamic expansion. Balanced clients will benefit from imbalanced clients using the proposed method of creating a collaboration mechanism of client layer-by-layer federation and intra-client cross-scale recursive fusion by Zhou F et al. (2024), Which increased the accuracy by 13.97% compared to existing methods[14].

Zhou, F (2021) paper designs a deep learning fault diagnosis network based on modular federation for bearing under different conditions. A deep neural network is used initially to extract features layer by layer, and then a dynamic routing algorithm is adopted. This paper used a new modular federated neural network based on FL[15]. Bharathi S et al. (2022) discussed challenges and future applications of federated learning by discussing data privacy, data protection, communication cost, system heterogeneity, Unreliable model upload, and future directions [16]. Wen et al. (2022) did a survey related to the practical application of learning and

bottlenecks. This paper reveals basic knowledge of FL, privacy and security protection mechanisms, communication overhead challenges, and heterogeneity problems in FL [17]. Lili et al. (2020) discussed the evolution of FL, realistic applications in IoT devices, and grounding applications in industrial engineering and healthcare. At the end of the paper, some frontier achievements are given about the future direction of FL[18]. There is a research gap in data optimization and standardization when implementing FL in rotating machinery. This paper addresses this as a novel system of FL in fault diagnosis. Figure 1 shows how to implement the federated learning approach in rotating machinery in the proposed methodology of this work.

3.1. Data Collection and Preprocessing from Rotating Machines

Data resources: Various sensors, like accelerometers, acoustic emission sensors, etc., have been mounted in the Rotating machines to collect the vibrational, acoustic, or temperature variational data. Data collection is the most important, which may lead to either accurate or less accurate predictions.

Storage of Local Data: After data collection, facility or device storage will store data locally by preserving data privacy.

3.2. Local Model Training Using Preprocessed Data

Development of model: Each facility on-site trains a local model using the collected data. This could be a machine learning model to detect faults, failures, or detection of anomalies.

Training methodologies: Training methodologies include Artificial neural networks, time series analysis, pattern detection, signal preprocessing, etc.

3.3. Federated Aggregation of Updates to a Central Server

Instead of sharing data, each facility uploads model updates to a central server. The central server aggregates the updates and refines the global model. The method for aggregation includes federated averaging, where the server averages from different facilities.

3.4. Global Model Update in the Central Server

The aggregated update improves the global model, and the updated global model is redistributed back to the facilities, which may allow the local model to benefit from collective knowledge served by the central server.

3.5. Evaluation of the Global Model and Iteration to Improve Diagnosis

The updated global model monitors the machinery based on its condition and potential faults. Also, the system can generate maintenance information. This process is iterative, and facilities continue to train their local models with new data from sensors.

Periodic updates will be sent to the central server to refine the global model, which adapts to new operating conditions and fault types. More data improves accuracy and reliability.

3.6. Phases of General FL Model

A general federated learning model will have many phases to standardize data to predict faults. In the initial stage, the data is shared with the central server after the fault prediction in the edge devices. Different edge device outputs will be merged using data merger for data preprocessing. During the data merging process, input from the participating devices was also uploaded to improve accuracy.

Data originality and origin validation will be done during the data processing, and the trained dataset will be generated. Then, the classification of records will be generated to create feature set mapping. The final process of FL data attributes and feature mapping will be created. Cluster optimization and data optimization are more important in the end process to have data standardization. The data standardization model in FL is given in Figure 2.

3.7. Challenges in Applying Federated Learning

The researchers face many challenges while applying federated learning in fault diagnosis of rotating machinery. The main challenges are given as follows.

1. Data transmission overhead: Efficient protocols related to the fault diagnosis data transmission are required to manage the communication between local facilities and the central server to minimize latency and bandwidth usage
2. Data heterogeneity: In the fault diagnosis of local facilities, various sensors and facilities may have

heterogeneous data. A proper technique is required to handle these heterogeneous data groups to improve the accuracy and consistency in predicting faults.

3. Model security and privacy: A potential update of each model security is needed to avoid possible attack risk or tampering. This is a highly challenging task in the fault diagnosis of rotating machinery. Otherwise, the prediction accuracy may be questionable and may give faulty results.
4. Compatibility: Systems must be compatible with fault diagnosis in different local facilities and for sensors like accelerometers, temperature sensors, pressure sensors, acoustic emission sensors, etc.
5. Amalgamation: Smooth integration between different local facilities to the central server is crucial in fault prediction and also needs to ensure accuracy in the functioning of the FL framework for condition monitoring
6. Law Agreement: The industries have policies and regulations for data handling and transmission. So, it is required to adhere to the company’s data protection laws, data security laws, and regulations related to the transmission of data from local facilities to servers and servers to local facilities during the continuous monitoring of real-time fault diagnosis.
7. Artificial Intelligence application: In the preprocessing of the FL framework, AI models have been used in the local facilities to make decisions related to faults on the site. The selection of an AI model for local facilities is a highly challenging job that may vary from one local facility to another.

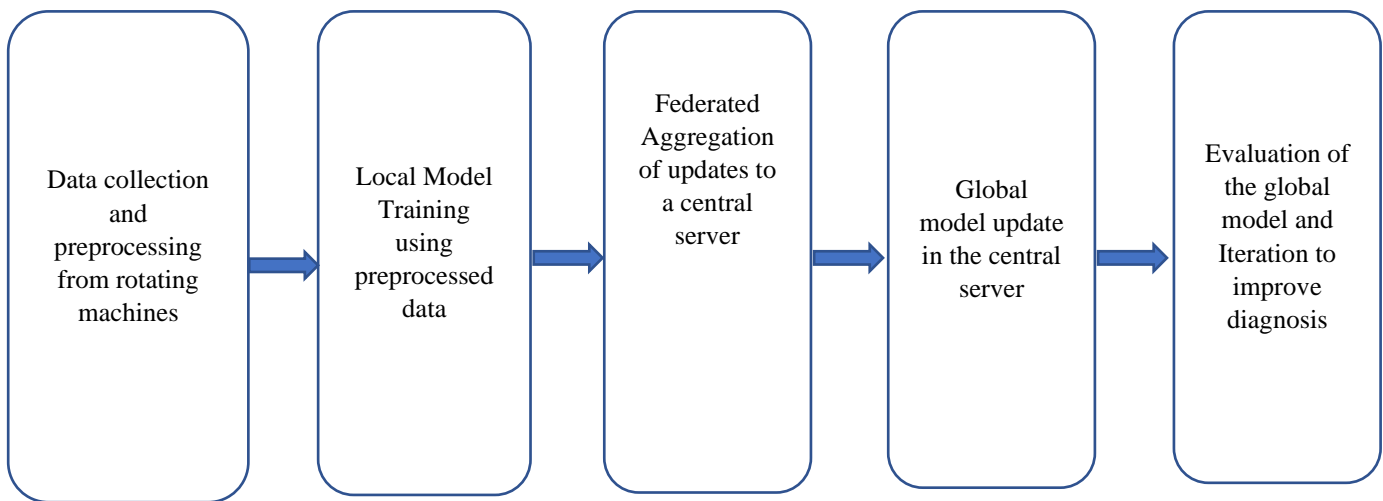


Fig. 1 The implementation process of federated learning

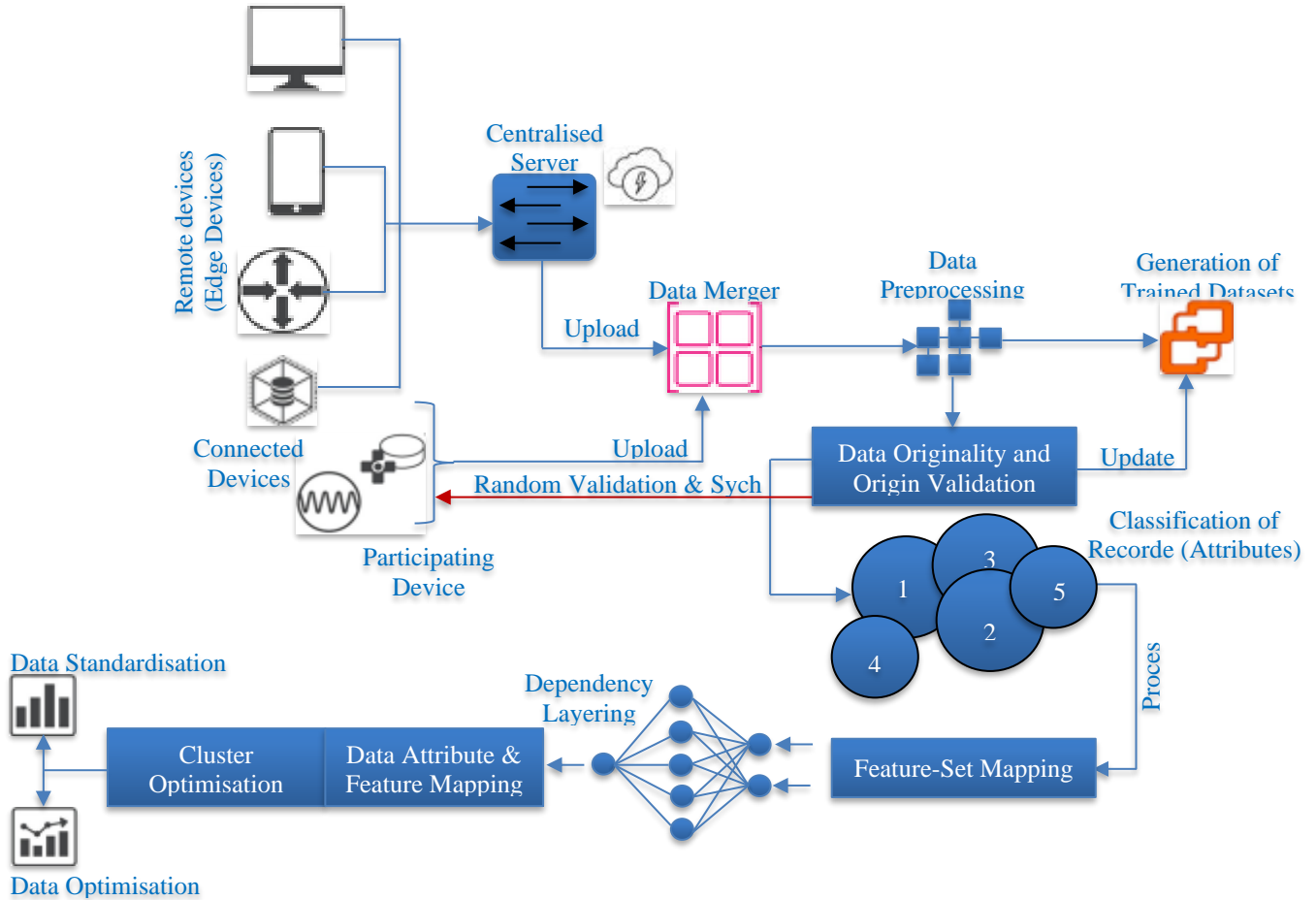


Fig. 2 Data standardization model in FL [19]

4. Model for Federated Learning-Based Fault Prediction System for Rotating Machinery

In the application of federated learning in fault diagnosis, object definition is critical before execution. We need to establish the goals of fault diagnosis and the types of faults that will be addressed. Then, an appropriate machine learning or deep learning model is required based on the data produced by sensors.

We must frame the FL model and its infrastructure in the third phase, including communication protocols. Then, local models will be trained, and the performance of the global model will be validated on the server. Monitoring system performance, updates, and data privacy must be addressed in these processes. The application of FL steps is given in Figure 3.

So, FL technology and its applications in fault diagnosis are based on data partitioning, architecture, the scale of the federation, and aggregation algorithms for generating the global model.

Data Partitioning contains three different types

1. Horizontal data partitioning
2. Vertical data partitioning
3. Hybrid data partitioning

This data partitioning determines the different types of FL, such as horizontal FL, vertical FL, and federated transfer learning.

Horizontal FL is for homogeneous datasets produced by different clients, which means the datasets have the same features. The data from each client is similar, and aggregating the data is much easier in this FL.

Vertical Data Partitioning is feature-based FL, in which data are heterogeneous. This is vertically partitioned to find overlapping data samples among clients, which are then used for training.

Hybrid Data Partitioning. Hybrid FL has partially overlapped feature spaces and sample spaces. This is applicable when all clients do not have all samples. Each client trains its local model on data with a subset of features. The server global model supports all the features.

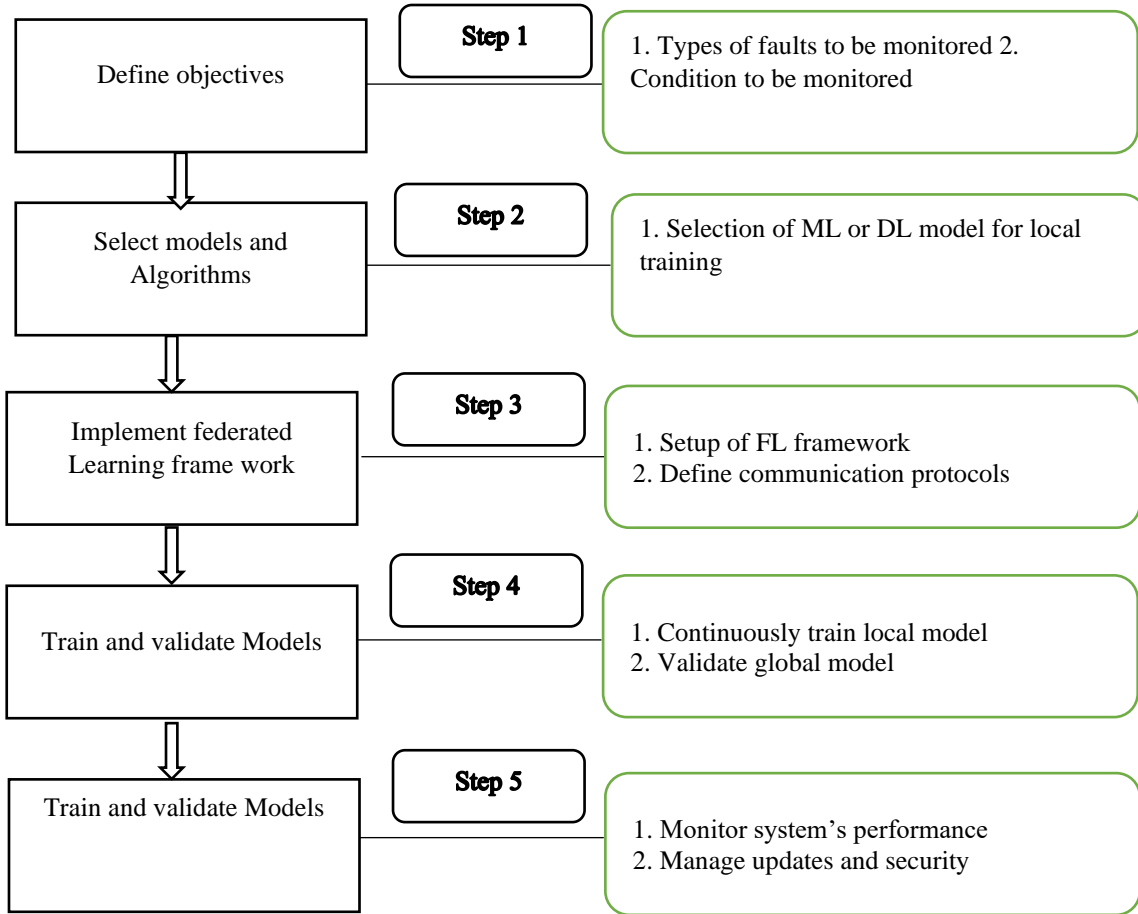


Fig. 3 Essential steps in condition monitoring of rotating machinery

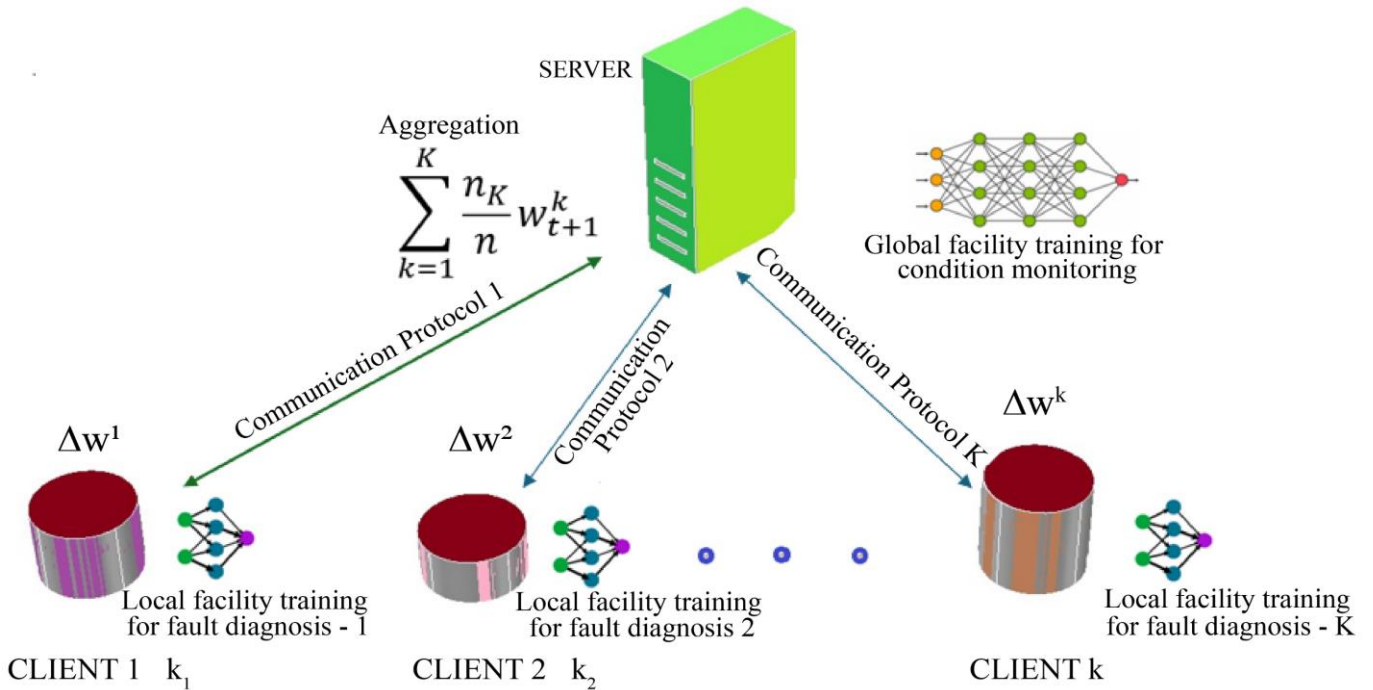


Fig. 4 Architecture of federated learning [20]

The architecture of federated learning is shown in Figure 4. It is essential to identify the number of data owners who want to use their data to train an AI model. In the Conventional practice of training an AI model is to combine all the data and apply $D = D_1 \cup D_2 \cup \dots \cup D_N$ to train a M_{SUM} model where the owners are $\{F_1, F_2, F_3, \dots, F_N\}$ of N data. The respective data is $\{D_1, D_2, \dots, D_N\}$. In a federated learning approach, the data owners will work together to train a model of M_{FED} , and no data owner F_i is required to share their data D_i with a centralized system. It should also be mentioned that the accuracy of V_{FED} is expected to be similar to that of M_{SUM} (or, more precisely, V_{SUM}). For the sake of argument, consider the case where there is a non-negative actual number (δ) If $V_{FED} - V_{SUM} < \delta$.

4.1. Mathematical Modeling of Federated Learning for Clients

X_1, X_2, \dots, X_k is the fault diagnosis data set of Client 1, Client..., and Client K , respectively. Assume that there is unbalancing in the clients. If so, $X_{1B}, X_{2B}, \dots, X_{kB}$ is the corresponding balanced dataset for the imbalance dataset. The FL model by the balanced dataset of each client can be formulated as

$$(w_{g,s+1,B}) = \sum_{k=1}^n p_{k,s,B} w_{g,s,B} \tag{1}$$

Where

$w_{g,s+1,B}$ – FL parameters using a balanced dataset of all clients
 $p_{k,s,B}$ – Model parameters of the k th client in the s th round of federation

$w_{k,s,B}$ – Corresponding aggregation weights of $p_{k,s,B}$

After establishing the dataset for each client, aggregation algorithms will be used in the central server to evaluate further fault diagnosis, as discussed in Section 5.

4.2. Centralized FL Architecture

Typically, the fault diagnosis FL system uses a centralized FL architecture in which the server has clients who collectively learn a global model without sharing data with another client. Only, the clients exchange weights (w) of their models with the server. So, optimization is described as

$$\min_{w \in R^d} f(w) \cong \frac{1}{m} \sum_{j=1}^m f_j(w) \tag{2}$$

Where $f_i: R^d$ in which R represents the loss function corresponding to client j . The centralized approach has some challenges related to delay due to bottlenecks, potential for system errors, and reliability concerns in creating a global model [21].

4.2.1. Scale of Federation

About the scale of training models and the number of local facilities, FL can be classified into two categories.

1. Cross-device FL
2. Cross-silo FL

Cross-device FL uses many edges like laptops, tabs, mobiles, IoT, etc. Each facility has a unique dataset for training. Cross-silo FL uses data from organizations or companies that are always available in the database.

Regarding the scale of training models and number of clients, FL can be classified into cross-device FL and cross-silo FL. Cross-device FL has many edges (i.e., IoT and mobile phones). Each client has a unique local row dataset for training local learning models. Also, FL applications with only a few edge devices form cross-silo FL. The number of clients is small, usually below a hundred. Clients in cross-silo FL are organizations or companies. Data from each organization in data silos are always available and typically have different example features.

4.3. Aggregation Algorithms

Aggregation algorithms in federated learning are important in updating the global model. Many aggregation approaches exist in building the aggregation algorithm in a federated learning environment.

The aggregation algorithm, which is highly suitable for fault diagnosis, are

- Average Aggregation
- Clipped Average Aggregation

4.3.1. Average Aggregation

Average aggregation is the initial approach and most commonly used aggregation. The server summarizes the model updates, parameters, and gradients in this. The set of participants denoted by “ N ” and their updates denoted by w is calculated using the following equation

$$w = (1 / N) * \sum_{i=1}^N w_i \tag{3}$$

4.3.2. Clipped Average Aggregation

In addition to the average aggregation steps, clipping the model updates to a predefined range has been added. This helps to reduce the impacts of outliers and malicious clients, similar to average aggregation. “ N ” denotes the set of participants, and w_i denotes their relative weights. $clip(x,c)$ is a function that clips the values of x to a range of ($-c$ to c) where c is the clipping threshold. Clipped aggregate is calculated as

$$w = (1 / N) * \sum_{i=1}^N clip(w_i, c) \tag{4}$$

A model of the aggregation algorithm is given in Figure 5.

Similar to the above aggregation, many aggregation methods exist according to data privacy, security, model change, weighted aggregate, parameter uncertainty, fraudulent model changes, communication efficiency, communication overhead, etc. The following are the types of aggregation to address the above features.

- Secure aggregation
- Differential privacy average aggregation
- Momentum aggregation
- Weighted aggregation
- Bayesian Aggregation
- Adversarial aggregation
- Quantization aggregation
- Hierarchical aggregation
- Personalized aggregation
- Ensemble aggregation

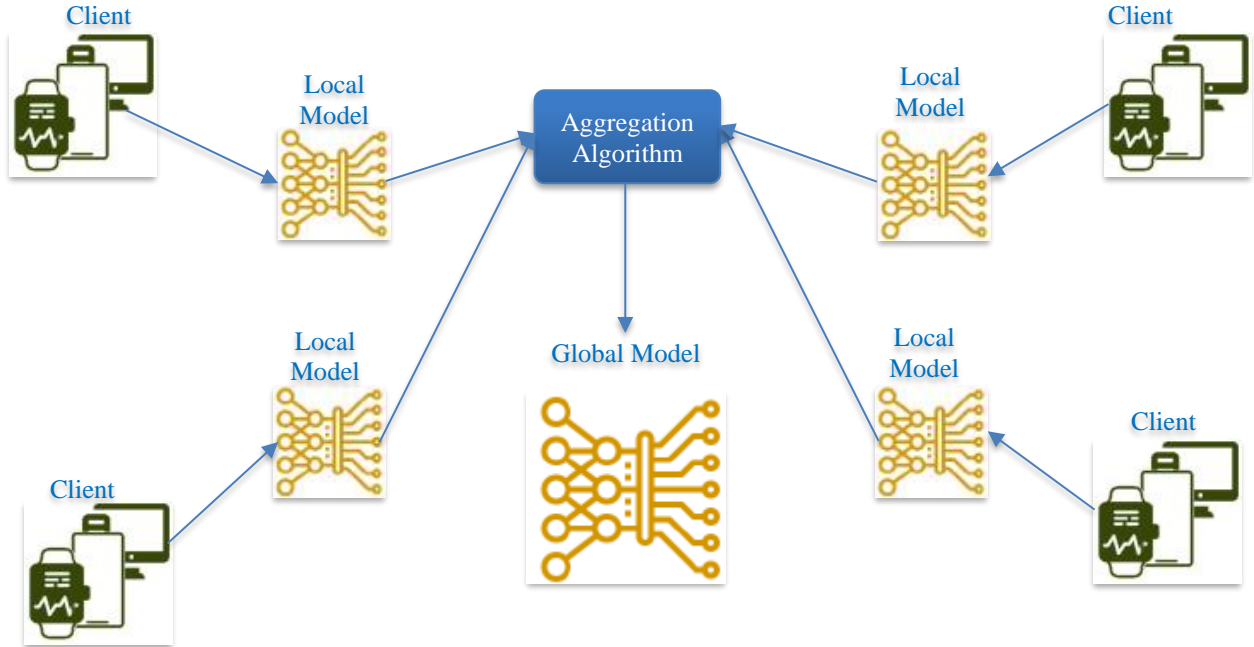


Fig. 5 Aggregation algorithms [22]

Average aggregation has many disadvantages, such as slow convergence, complex tuning, high communication and computational costs, client heterogeneity, scalability, and more. These problems need to be solved by developing algorithms to manage several issues.

5. Proposed Federated Learning Model for Fault Diagnosis of Rotating Machinery - FLOACOS Method

Systematic data handling is needed to improve the accuracy in predicting a fault in rotating machinery located in different production systems. Figure 6 shows the federated learning-based futuristic fault diagnosis and standardization model. This proposed FL model for fault diagnosis contains 4 phases: fault data origin phase, fault data acquisition phase, fault data classification phase, and fault data optimization.

5.1. Fault Data Origin Phase

The fault diagnosis of rotating machinery assumes that the clients are spread from 1 to k in various sites, which means fault data originates from different locations. So, multiple machines will give fault data, which is the data source for creating a dataset. The fault data from the data source will be collected using various sensors like accelerometers, acoustic emission sensors, etc. Since local facilities provide continuous

data, big data will be collected. To store all big data, a cloud concept should be introduced to manage the database. Different user interfaces can be used to improve the data collection speed in this data collection.

5.2. Fault Data Acquisition Phase

The second phase in the proposed model is the fault data acquisition phase, in which data preprocessing will be done to collect the required data for further processing. In this preprocessing stage, data cleaning and filtering concepts will eliminate the unwanted data or noise attached to faulty data. In fault diagnosis, vibration or sound data will be collected in which noise will be added as a disturbance.

The noise signal is not required for further processing, so data filtering is needed to eliminate noise signals. This preprocessing aim is to extract the features of the fault signal to do feature mapping. AI models can extract significant features to handle the fault data. In feature extraction processing, multiple features were extracted from the original vibration signals, followed by the dimension of the high feature set. Feature extraction techniques will vary for time domain and frequency domain fault signals. Frequency domain features are generally more consistent in detecting damage than time domain parameters [23].

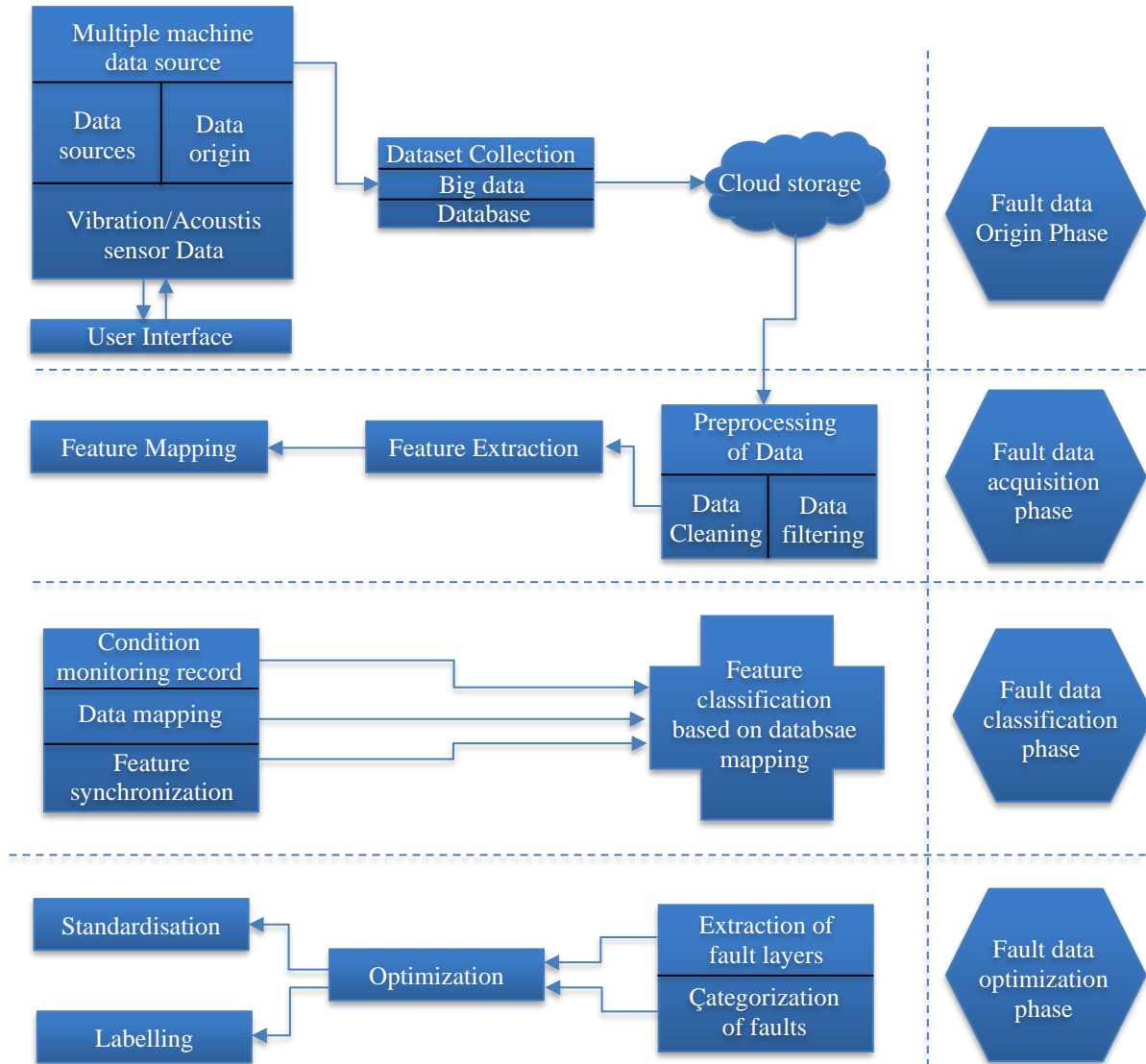


Fig. 6 Block diagram of proposed federated learning based futuristic fault diagnosis model

5.3. Fault Classification Phase

In this phase, the feature mapping data is transferred to database mapping. A condition monitoring record will be maintained in the database. For example, if the fault data is imbalanced, balanced data will be in the condition monitoring record, which is used for comparison and feature synchronization. The observer or diagnosing software can only obtain partial information regarding the system state at irregular intervals determined by specific synchronizing conditions between the system and the observer or diagnosis [24]. Condition monitoring records, data mapping, and feature synchronization are used for feature classification, which will be used to extract fault layers.

5.4. Fault Data Optimization

After feature classification from the fault data collected from different local facilities, fault layers will be extracted. Then, the categorization of faults will be demonstrated.

Various faults may come from rotating machines, such as misalignment, crack, unbalancing, bearing faults, gear faults, cavitation, etc. So, categorization is more critical to fixing future faults by the industrial sector’s maintenance team. Also, it is required to identify the faults and the respective local facility that has produced the fault dataset. After extracting fault layers and categorizing faults, dataset optimization is needed to standardize fault detection. Usually, data produced by the local facilities will be raw, requiring meaningful information labeling. Labeled data will be carefully annotated with meaningful tags or labels that classify the data’s elements or outcomes. Fault data standardization is a data processing workflow that converts the structure of different fault datasets into one standard data format. Standardized data is the data from various sources that has been transformed into a consistent, standards-based format. The data standardization process involves harmonizing data so that the different fault datasets relating to the same terms and format allow for

comparison with faults. Data standardization is required to process aggregation, in which the accuracy of the fault prediction will be increased on the server.

5.5. Model Structure

The first part of the model execution is data acquisition from the vibration sensor (accelerometer), and then the signal will undergo the signal conditioning process. After acquiring the signal using LabVIEW Software, deep learning methods such as Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network, and Long Short Term Memory (LSTM) will be used under the proposed FL framework. In the training of the deep learning model, datasets from different clients will be collected. Using the FL OACOS method, the dataset has been standardized and compared with the best-performing deep learning model.

5.6. Algorithm of Proposed Model

Algorithm of the Proposed Federated Learning Model is as follows:

- Step 1: Server Initialize the model parameter
- Step 2: Send server instructions to clients
- Step 3: The client performs data optimization and standardization using the FLOACOS method
- Step 4: Used deep learning method, and model weights will be uploaded to the server
- Step 5: Unsupervised clustering by server using client’s representation
- Step 6: Aggregation process to the corresponding clients

6. Experimental Setup

To validate the effectiveness of the proposed OACOS FL combined with DL algorithms of the fault diagnosis model, unbalancing data from the spectra machine fault simulator has been used. The Spectra Quest Machinery Fault Simulator (SQMFS) simulates rotary shafts’ balanced and imbalanced conditions. The Machine Fault Simulator (MFS) is a highly versatile and modular test platform designed to facilitate the study and simulation of various mechanical faults in rotating machinery. The methodology of data collection using a Machine fault simulator is shown in Figure 7. In the machine fault simulator, unbalancing was formed using different weights inserted in the disc mounted on the shaft. An accelerometer is placed vertically to collect the datasets using the data acquisition card NI CDAQ- 9172 and the signal conditioning device NI 9234. This close-up image of the MFS (Figure 7) rotor provides a detailed view of the critical components. We can see the motor-driven shaft, adjustable disks with tapped holes for attaching weights, and the accelerometer mounted to capture vibration data. The golden disk on the shaft is one of the adjustable disks for introducing imbalance (Figure 3). This configuration allows for precise control over the introduction of faults and the measurement of resulting vibrations. Lab VIEW software has been used for signal collection in the time and frequency domains.

In this experimental work, MFS is considered a client and different data sets were collected for unbalancing by generating different readings by considering other clients. The data received from the machine fault simulator were analyzed using multilayer perceptron, CNN, RNN, and LSTM. The data collection model in the proposed methodology is given in Figure 8.

The experimental analysis focused on two primary operating conditions of the rotary shaft:

1. Balanced Condition
2. Imbalanced Condition

These conditions were carefully established to study the effects of imbalance on the dynamic behavior of rotating machinery, thereby providing critical insights for developing diagnostic tools. In data collection, the motor was operated at 30 HZ and with a gear or shaft frequency of 216 HZ.



Fig. 7 Machine fault simulator to collect the unbalancing dataset

The evaluated model’s performance using the original dataset, focusing on critical metrics like accuracy, precision, recall, and F1-score, is shown in Table 1 for client 1. Table 2 Shows the Accuracy in the prediction of unbalancing in the four clients using different algorithms.

As per the table, RNN and LSTM are performing poorly compared to MLP and CNN. MLP and CNN performed better and received 90% and 95% efficiency. Prediction accuracy in a one-point scale is given for various DL algorithms in Figure 9.

Table 1. Model performance for standard dataset

	Model	Accuracy	Precision	Recall	F1-Score
0	MLP	0.900	0.917391	0.900	0.899499
1	CNN	0.950	0.954762	0.950	0.950000
2	RNN	0.575	0.646094	0.575	0.533908
3	LSTM	0.450	0.453419	0.450	0.433120

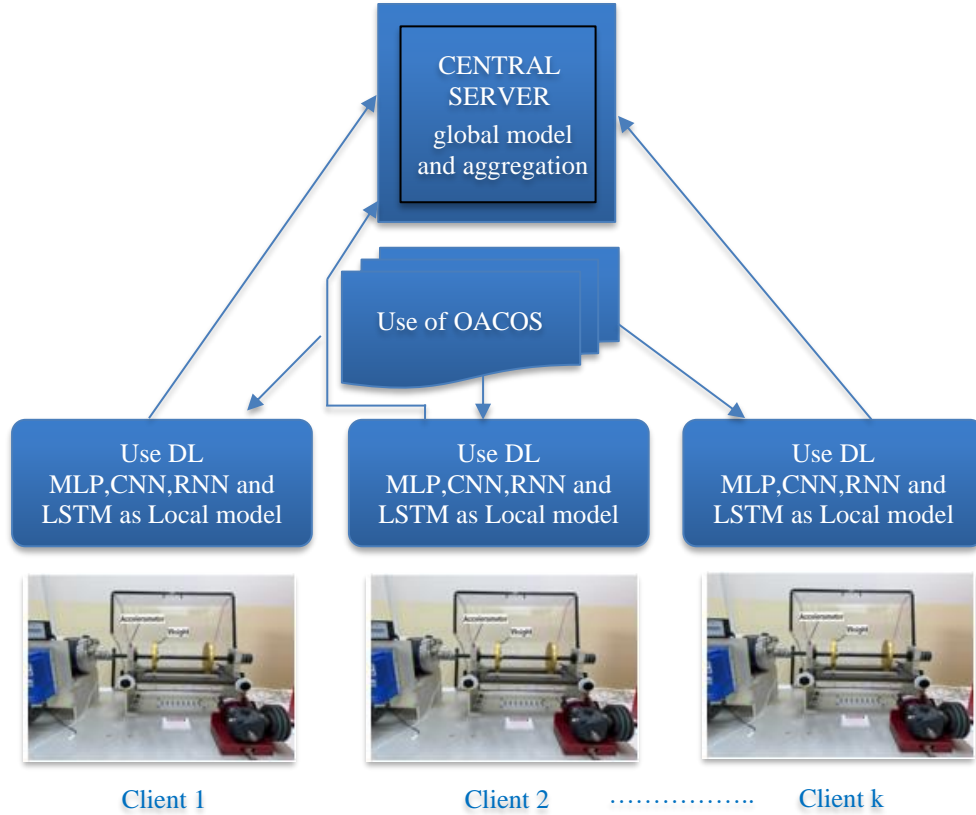


Fig. 8 Data collection model from clients

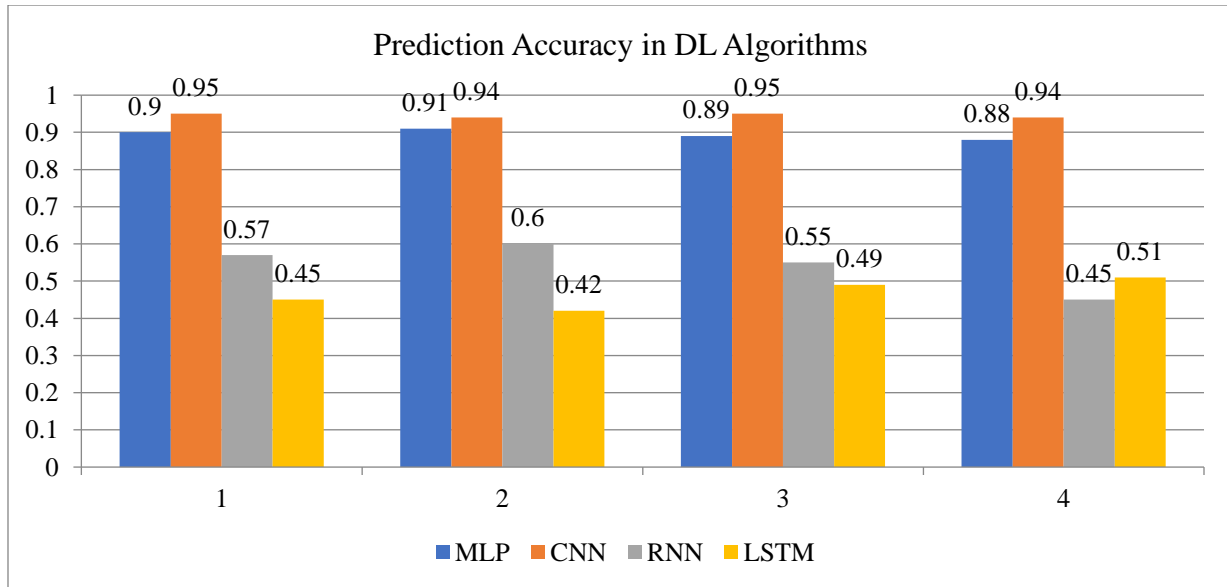


Fig. 9 Prediction accuracy in various DL algorithms

FLOACOS has been applied to standardize the data before execution using deep learning to improve the efficiency of fault prediction in local facilities. Since RNN and LSTM are not performing well in fault prediction, these DL algorithms are not taken for further progress. MLP and CNN are integrated with FLOACOS and then predicted accuracy.

OACOS-modelled MLP increased the accuracy by 3.14%, and CNN with OACOS increased the accuracy by 3.56%. So, overall accuracy came in CNN as 98.56%. Table 3 shows the accuracy of client 1 with and without FLOACOS. A comparison chart after applying the proposed methodology of FLOACOS is given in Figure 10.

Table 2. Accuracy in prediction of unbalancing in the four clients using different DL algorithms

DL Algorithm	Client 1	Client 2	Client 3	Client 4
MLP	0.90±0.34	0.91±0.05	0.89±1.10	0.88±1.32
CNN	0.95 ± 0.10	0.94±0.40	0.95±1.04	0.94±0.80
RNN	0.57±0.5	0.60±0.15	0.55±0.34	0.45±0.34
LSTM	0.45±0.0	0.42±0.23	0.49±0.21	0.51±0.0

Table 3. Comparison of the application of DL in the local model without proposed FLOACOS and with FLOACOS

DL Algorithm	Client 1
MLP	0.90±0.34
CNN	0.95 ± 0.50
MLP with OACOS	0.93±0.48
CNN with OACOS	0.98 ±0.66

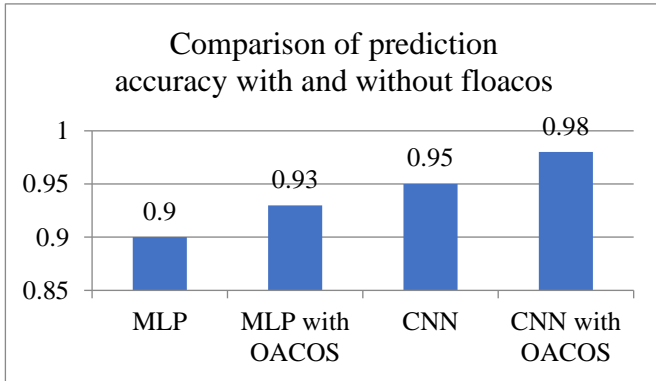


Fig. 10 Comparison chart after applying the proposed methodology

Many challenges related to data handling, privacy, and security have been faced in the execution, data preprocessing, data standardization, application of DL, and selection of communication protocols. The choice of the aggregation model is also quite challenging when implementing FL in fault diagnosis. There may be a more significant challenge in applying the algorithm in a real-time environment than in a machine fault simulator application. In the real-time data set, hybrid algorithms combined with FLOACOS may help maintain accuracy in fault diagnosis and overcome challenges.

7. Conclusion and Future Research Direction of FL in Fault Diagnosis

Overall, federated learning offers a promising approach for enhancing fault diagnosis in rotating machinery by leveraging centralized or decentralized data while maintaining privacy and improving diagnostic accuracy. Organizations can achieve a robust, privacy-preserving solution that leverages collective intelligence to enhance predictive maintenance and

operational efficiency by utilizing federated learning for condition monitoring of rotating machinery. Due to the advancement of digital technologies and wireless sensor networks, maintenance engineering has been shifted to a new paradigm in fault diagnosis. Many researchers are working on edge computing and 5G networks using FL features. Nguyen T.D. et al. used FL to enhance attack detection accuracy and communication efficiency in a smart home wireless IoT system [25]. FL can be used in medical diagnosis, healthcare, aerospace, marine systems, manufacturing, process operations maintenance, and office automation [26].

The future direction of FL will extend to extreme communication schemes, communication reduction, and the Pareto frontier, focusing on novel models of asynchrony, heterogeneity diagnosis, granular privacy constraints, beyond supervised learning, productionizing FL, benchmarking of FL approach, etc. In this research, a novel idea for data standardization is proposed and implemented in FL based on Fault diagnosis in rotating machinery. The proposed methodology is applied in the best unbalancing fault prediction algorithms of MLP and CNN. The result shows that the improvement in accuracy is 3.5% in both cases. So, FL will provide a safe model for fault diagnosis when data is shared. It will tackle the issue of data privacy when more local facilities are used in production systems [27]. Overall. There are many advantages and disadvantages to the FL-based fault diagnosis [28]. The proposed method will overcome many disadvantages the previous researchers listed and provide enhanced results in the condition monitoring of rotating machinery.

The future research direction of FL will use a hybrid algorithm to predict the heterogeneous group of datasets. In this work, unbalancing is considered the fault of all four clients. In real-time, there may be some different faults that will produce heterogeneity in nature. To improve the accuracy of FL, future researchers can move toward a new algorithm that will overcome the heterogeneity.

Author Contributions

Conceptualization, KV, and RA.; Literature collection, KV, RA, and S.RK.A.V; Formal analysis, RA, KV, and S.RK.A.V; Funding acquisition, RA, KV, and KPR; Investigation, RA, KV, and KPR; Methodology, RA, KPR, KV, and S.RK.A.V; Validation, RA, and KV; Writing an original draft, RA; Review and editing, RA, KPR, KV, and S.RK.A.V. All authors have read and agreed to the published version of the manuscript

Funding

This work was performed as part of the research work titled” Deep learning based real-time drill condition monitoring using acoustic emission sensor,” funded by MOHERI, OMAN, grant number BFP/RGP/EI/22/433

Acknowledgments

The authors express their sincere thanks to the research council, Sultanate of Oman, for funding the research work and

also grateful to their affiliated university, the National University of Science and Technology, Oman, for their in-kind contribution and support.

References

- [1] Manan Mehta et al., "A Federated Learning Approach to Mixed Fault Diagnosis In Rotating Machinery," *Journal of Manufacturing Systems*, vol. 68, pp. 687-694, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [2] Weihua Li et al., "Clustering Federated Learning for Bearing Fault Diagnosis in Aerospace Applications with a Self-Attention Mechanism," *Aerospace*, vol. 9, no. 9, pp. 1-19, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Yaguo Lei et al., "Applications of Machine Learning to Machine Fault Diagnosis: A Review and Roadmap," *Mechanical Systems and Signal Processing*, vol. 138, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [4] Ji Liu et al., "From Distributed Machine Learning to Federated Learning: A Survey," *Knowledge and Information Systems*, vol. 64, pp. 885-917, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] Yaguo Lei, and Ming J. Zuo, "Gear Crack Level Identification Based on Weighted K Nearest Neighbor Classification Algorithm," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1535-1547, 2009. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] B. Samanta, "Gear Fault Detection Using Artificial Neural Networks and Support Vector Machines with Genetic Algorithms," *Mechanical Systems and Signal Processing*, vol. 18, no. 3, pp. 625-644, 2004. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Ruonan Liu et al., "Artificial Intelligence for Fault Diagnosis of Rotating Machinery: A Review," *Mechanical Systems and Signal Processing*, vol. 108, pp. 33-47, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] Rui Zhao et al., "Deep Learning and Its Applications to Machine Health Monitoring," *Mechanical Systems and Signal Processing*, vol. 115, pp. 213-237, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Qizhao Wang et al., "Efficient Federated Learning for Fault Diagnosis in Industrial Cloud-Edge Computing," *Computing*, vol. 103, pp. 2319-2337, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Zexi Yao et al., "DiagNet: Machine Fault Diagnosis Using Federated Transfer Learning in Low Data Regimes," *AAAI 2022 FL Workshop*, 2022. [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Wei Zhang et al., "Federated Learning for Machinery Fault Diagnosis with Dynamic Validation and Self-Supervision," *Knowledge-Based Systems*, vol. 213, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Manar Abdelmaksoud et al., "Convolutional-Neural-Network-Based Multi-Signals Fault Diagnosis of Induction Motor Using Single and Multi-Channels Datasets," *Alexandria Engineering Journal*, vol. 73, pp. 231-248, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Hao Lu et al., "Federated Learning with Uncertainty-Based Client Clustering for Fleet-Wide Fault Diagnosis," *Mechanical Systems and Signal Processing*, vol. 210, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [14] Funa Zhou et al., "Fault Diagnosis Based on Federated Learning Driven by Dynamic Expansion for Model Layers of Imbalanced Client," *Expert Systems with Applications*, vol. 238, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] Funa Zhou, Zhiqiang Zhang, and Sijie Li, "Research on Federated Learning Method for Fault Diagnosis in Multiple Working Conditions," *Complex Engineering Systems*, vol. 1, no. 7, pp. 1-17, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [16] Subrato Bharati et al., "Federated Learning: Applications, Challenges and Future Directions," *arXiv*, pp. 1-35, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] Jie Wen et al., "A Survey on Federated Learning: Challenges and Applications," *International Journal of Machine Learning and Cybernetics*, vol. 14, pp. 513-535, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] Li Li et al., "A Review of Applications in Federated Learning," *Computers & Industrial Engineering*, vol. 149, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] Afifa Salsabil Fathima et al., "Federated Learning Based Futuristic Biomedical Bigdata Analysis and Standardization," *PLoS ONE*, vol. 18, no. 10, pp. 1-16, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Krista Rizman Žalik, and Mitja Žalik, "A Review of Federated Learning in Agriculture," *Sensors*, vol. 23, no. 23, pp. 1-20, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Funa Zhou et al., "Federated Learning Based Fault Diagnosis Driven by Intra-Client Imbalance Degree," *Entropy*, vol. 25, no. 4, pp. 1-18, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Mohammad Moshawrab et al., "Reviewing Federated Learning Aggregation Algorithms; Strategies, Contributions, Limitations and Future Perspectives," *Electronics*, vol. 12, no. 10, pp. 1-35, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Hongyu Yang, Joseph Mathew, and Lin Ma, "Vibration Feature Extraction Techniques for Fault Diagnosis of Rotating Machinery: A Literature Survey," *Asia-Pacific Vibration Conference*, pp. 1-7, 2003. [[Google Scholar](#)] [[Publisher Link](#)]
- [24] E. Athanasopoulou, and C.N. Hadjicostis, "Synchronization-Based Fault Detection in Discrete Event Systems," *2004 43rd IEEE Conference on Decision and Control*, Nassau, Bahamas, vol. 1, pp. 57-62, 2004. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]

- [25] Thien Duc Nguyen et al., “D³IoT: A Federated Self-learning Anomaly Detection System for IoT,” *arXiv*, pp. 1-12, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [26] Yan Song, and Peng Liu, “Federated Domain Generalization for Intelligent Fault Diagnosis Based on Pseudo-Siamese Network And Robust Global Model Aggregation,” *International Journal of Machine Learning and Cybernetics*, vol. 15, pp. 685-696, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Tarek Berghout et al., “Federated Learning for Condition Monitoring of Industrial Processes: A Review on Fault Diagnosis Methods, Challenges, and Prospects,” *Electronics*, vol. 12, no. 1, pp. 1-20, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [28] Shiyi Jiang et al., “Federated Clustered Multi-Domain Learning for Health Monitoring,” *Scientific Reports*, vol. 14, no. 1, pp. 1-12, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]