International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

Communication-Efficient Federated Learning for IIoT Bearing Fault Diagnosis Using Multi-Head Deep Q-Network-Based Intelligent Client Selection

[1]Bhupathirao Lakinana, [2]Kalyana Chakravarthy Chilukuri

[1]Assistant Professor, Department of Computer Science & Engineering - Artificial Intelligence, Vignan's Institute of Information Technology (Autonomous), Visakhapatnam, Andhra Pradesh, India.
[2]Professor, Department of Computer Science & Engineering, MVGR College of Engineering (Autonomous), Vizianagaram, Andhra Pradesh, India.
[1]bhupathilakinana@gmail.com, ORCID ID: 0000-0002-0100-2177
[2]kch.chilukuri@gmail.com, ORCID ID: 0000-0002-8137-7074

**Corresponding Author: Bhupathirao Lakinana*

Abstract: Federated learning (FL), which allows distributed training across several industrial machines without exchanging raw vibration data, has emerged as a promising paradigm for privacy-preserving bearing fault diagnosis in Industrial Internet of Things (IIoT) scenarios. Nevertheless, current FL approaches for fault diagnosis face two significant limitations. First, uniform client participation in each communication round results in needless and prohibitive communication overhead. Second, static or random client selection strategies are indifferent to the informativeness of individual client updates under highly non-IID data distributions. This paper proposes a novel Multi-Head Deep Q-Network (Multi-Head DQN)-based client selection framework for communication-efficient federated learning applied to bearing fault diagnosis to overcome these limitations. In the proposed framework, the Multi-Head DQN agent jointly determines the optimal number of clients, *k*, to select in each round from a predefined range, and the identities of the most informative clients. To accurately replicate heterogeneous IIoT deployments in the real world, a stringent non-IID partitioning approach is implemented across 12 federated clients, each of which is allocated precisely two of four failure classes (Normal, Inner Race, Outer Race, and Ball). The Flower federated learning framework is used to compare six FL configurations: FedAvg, FedProx, FedAvg+FedRandom, FedProx+FedRandom, FedAvg+FedDQN, and FedProx+FedDQN. All experiments were conducted on the CWRU bearing fault dataset. Experimental results show that the proposed FedProx+FedDQN approach reduces cumulative communication costs by 19.15% compared with the all-client FedAvg baseline while achieving a global fault classification accuracy of 94.05%, which is comparable to the full-participation baselines (FedAvg and FedProx). The DQN-based selection prioritizes informative client updates, leading to faster convergence and lower communication overhead. These findings confirm that intelligent client selection is an effective strategy for improving communication efficiency in federated learning for industrial fault diagnosis applications.

Keywords: Federated learning, bearing fault diagnosis, Industrial Internet of Things, deep Q-network, client selection, communication efficiency, non-IID data, convolutional neural network, reinforcement learning.

1 INTRODUCTION

In contemporary production settings, condition monitoring and predictive maintenance have changed significantly due to the explosive growth of the IIoT. A considerable amount of unscheduled industrial downtime is caused by rolling element bearings, which are among the most important components and most prone to failure in rotating machinery [1]. Therefore, in IIoT-driven Industry 4.0 systems, timely and accurate bearing fault identification is crucial for guaranteeing operational reliability, cutting maintenance costs, and preventing catastrophic equipment failures. Many machine learning and deep learning algorithms have been developed for early bearing fault detection. Machine learning methods require manual feature engineering, whereas deep learning methods, in particular one-dimensional convolutional neural networks (1D-CNNs), have shown exceptional performance in automatically extracting features from raw vibration data, significantly surpassing manually designed feature engineering techniques [2][3][4].

Traditional deep learning methods for diagnosing bearing faults rely on centralized data collection, which involves sending vibration signals from all industrial equipment to a central server for collaborative model training. However, this paradigm faces significant practical challenges in IIoT deployments: the enormous amount of continuous sensor data makes centralized aggregation expensive in terms of communication bandwidth; data privacy regulations in industrial settings increasingly restrict cross-enterprise data sharing; and raw vibration data from individual machines may constitute commercially sensitive operational intelligence [5]. These constraints have motivated a growing body of research on distributed and privacy-preserving learning paradigms for IIoT fault diagnostics.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

First presented by McMahan et al. [6], federated learning provides a feasible solution to these problems by allowing several clients to cooperatively train a common global model without sharing raw data. The central server uses the Federated Averaging (FedAvg) method to aggregate the model parameters that each client in the FL paradigm updates locally on its private data. FL is particularly useful for cross-enterprise and cross-factory IIoT deployments because it allows each industrial machine to contribute to a global fault classifier for bearing fault diagnosis without disclosing its private vibration data.

Despite its potential, FL in IIoT environments faces two interrelated fundamental challenges. The first is statistical heterogeneity (non-IID data): in actual industrial deployments, many machines operate under varying operating conditions, wear states, and fault modes, leading to client data distributions that are far from independent and identically distributed (IID) [7]. Under highly non-IID data distributions, where each machine contains only a portion of the potential fault classes, FedAvg exhibits degraded convergence and reduced global model accuracy due to the heterogeneity of local client updates. FedProx [8] has been proposed to address statistical heterogeneity and shows considerably more stable convergence than FedAvg in heterogeneous environments. It does this by augmenting the local training objective with a proximal regularization term to restrict divergence from the global model. The server must broadcast the global model to every client during each FL communication round and receive updated parameters in return, which presents the second significant challenge and results in significant overall communication overhead [9]. This bottleneck is exacerbated in wireless IIoT situations with constrained bandwidth. Therefore, a crucial research goal is to lower the number of participating clients per round while preserving model accuracy.

Model compression and client selection are two main approaches for improving communication efficiency in FL. A particularly serious limitation under non-IID distributions, where some clients may provide redundant or conflicting updates, is that random client selection, in which a fixed fraction of clients participate each round, lowers communication costs but ignores the quality and informativeness of individual updates [10]. Consequently, client selection based on reinforcement learning (RL) has become an effective complement that dynamically identifies the most informative clients. Wang et al. [11] proposed fast aggregation and valuation of client contributions for heterogeneous FL (FAVOR), an experience-driven DQN framework for client selection that balances non-IID bias and speeds up convergence. Xu et al. [12] proposed FedCG, which combines gradient compression and diversity-aware client selection for communication efficiency in heterogeneous edge systems. Although FAVOR [11] and FedCG [12] propose client selection strategies in federated learning, they focus only on selecting a subset of clients, and their methods were not evaluated on bearing fault diagnosis datasets. In contrast, the proposed method jointly determines both the optimal number of participating clients and the specific clients to be selected in each communication round for bearing fault classification in IIoT environments. This work applies a joint client selection strategy for bearing fault diagnosis in IIoT environments by simultaneously determining both the number of participating clients and their identities.

This paper proposes a Multi-Head DQN-based client selection framework that jointly determines the number of participating clients and their identities. To objectively assess model performance under stringent non-IID conditions, an unbiased server-side global test set constructed using CWRU data is utilized, together with a physically consistent reward function based on actual bidirectional communication cost. Through a comprehensive assessment of six federated learning configurations, the proposed framework is thoroughly validated in the Flower FL environment, highlighting the contributions of client selection, local training methods, and their combined consequences. The remainder of this paper is organized as follows. Related research on FL for fault diagnosis, non-IID FL optimization, and RL-based client selection is reviewed in Section 2. The proposed methodology, including the dataset, non-IID partitioning, CNN architecture, FL framework, communication cost model, and Multi-Head DQN design, is explained in Section 3. The experimental results and findings are shown in Section 4. The paper is concluded in Section 5.

2 LITERATURE REVIEW

This section reviews three closely linked research areas: federated learning for diagnosing bearing faults, non-IID data issues and strategies to improve performance in FL, and client selection techniques for communication-efficient FL.

2.1. Federated Learning for Bearing Fault Diagnosis

Sun *et al.* [13] demonstrated that FedAvg-based aggregation degrades when local fault class distributions are unbalanced. FedAlign is a data-free knowledge distillation technique that aligns global and local models to handle data heterogeneity in FL-based machine fault detection. Zhou *et al.* [14] proposed an imbalance-degree-guided federation technique that improves fault detection accuracy under heterogeneous client data distributions by adjusting aggregation weights depending on intra-client class imbalance. Sun *et al.* [15] proposed a federated adversarial technique motivated by fault information discrepancy to reduce negative transfer in FL-based fault detection, where poor local models impair the performance of global models. Ke et al. [16] proposed a prototype contrast federated transfer learning framework for multi-condition rolling bearing fault diagnosis, improving diagnostic performance under varying operating conditions. Although these methods improve bearing fault diagnosis under heterogeneous client data, they rely on effective mitigation of statistical heterogeneity (non-IID data), which remains one of the fundamental challenges in federated learning. Therefore, the following section reviews representative approaches for addressing non-IID data distributions.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

2.2. Non-IID Data Challenges and Mitigation in Federated Learning

Ma *et al.* [17] examined various federated learning approaches for managing non-IID data and categorized them into data sharing, client grouping, aggregate re-weighting, and personalization strategies. Building on this categorization, recent studies have shown that statistical heterogeneity, or the distribution of non-IID data among FL clients, is a major barrier to achieving fast and accurate federated learning in practice [18]. Local model updates may point in opposing directions when local data distributions deviate significantly from the global distribution, leading to client drift, a condition that severely reduces convergence speed and final model correctness. To address this issue, S. P. Karimireddy et al. [19] theoretically analyzed client drift through SCAFFOLD and proved that heterogeneous (non-IID) client data causes unstable convergence in FedAvg. This issue is particularly severe in IIoT-based fault diagnosis, where different machines operate under varying conditions and naturally contain only a subset of fault categories, resulting in highly non-IID client data.

Li et al. [8] further demonstrated that FedProx prevents local updates from drastically deviating from the global model by incorporating a proximal regularization term in the local objective function to handle client drift. FedProx outperforms FedAvg in extremely heterogeneous circumstances in terms of test accuracy and convergence stability. Although algorithms such as SCAFFOLD and FedProx effectively mitigate the optimization challenges caused by non-IID data, they typically assume that all or a large fraction of clients participate in every communication round. In practical IIoT systems, this assumption incurs significant communication overhead, motivating efficient client selection strategies.

2.3. Client Selection for Communication-Efficient Federated Learning

One well-researched issue in communication-efficient FL is reducing the number of clients participating in each FL round while maintaining model accuracy. FedCS, a technique that chooses clients with sufficient processing and communication power to complete training on schedule, was proposed by Nishio and Yonetani [20]. It minimizes delays brought on by slow clients, but it ignores variations in how clients distribute their data. Power-of-choice selection was proposed by Cho et al. [21], who showed that under extreme non-IID conditions, biased selection of clients with higher local losses accelerates convergence. However, excessive selection bias may degrade the final model accuracy. To balance exploration and exploitation in client selection, bandit-based techniques [22] employ multi-armed bandit algorithms, which provide theoretical performance guarantees but assume a fixed number of clients per round. Although these approaches improve communication efficiency and convergence, they rely on predefined heuristics or simplified assumptions, such as fixed client participation or manually designed selection policies. These limitations motivate adaptive learning-based client selection strategies.

To overcome the limitations of heuristic and bandit-based methods, RL has become a suitable paradigm for adaptive client selection. Unlike conventional methods, RL continuously learns an optimal client selection policy from interactions with the federated learning environment without requiring explicit knowledge of client data distributions. Among RL approaches, DQN-based client selection methods have attracted attention because they can learn client selection policies in dynamic non-IID environments. The first DQN-based client selection framework for non-IID FL, FAVOR, was proposed by Wang et al. [11]. It selects clients to counteract non-IID bias by using client model weights as a proxy for local data distribution. The number of clients chosen per round, k, is a fixed hyperparameter that is not dynamically adjusted, even though FAVOR showed notable gains in convergence performance. To increase action diversity, the FedRLCS framework [23] expanded FAVOR with a double DQN DDQN and top-p sampling; nonetheless, it treats k as fixed.

Wang [24] proposed FedCcs, a cluster-based client selection algorithm for Non-IID federated learning that improves convergence and test accuracy through computationally efficient client clustering. Despite the success of existing RL-based client selection methods, they treat the number of participating clients and client identity as two independent problems. Most methods assume a fixed number of selected clients and only optimize which clients participate in each round. Consequently, they cannot dynamically balance communication cost and model performance when client availability and data characteristics change. In bearing fault diagnosis, client devices often collect vibration signals under different operating conditions, resulting in highly heterogeneous and non-IID data distributions. Therefore, adaptive client selection is essential for achieving efficient communication while maintaining diagnosis accuracy. To address this limitation, this paper proposes a multi-head DQN framework that jointly determines how many clients should participate and which clients should be selected in every communication round. The dual-head architecture enables simultaneous optimization of communication efficiency and model accuracy, making it particularly suitable for bearing fault classification under heterogeneous and non-IID federated environments.

3 Proposed Methodology

This section presents the proposed communication-efficient federated learning approach for IIoT bearing fault diagnosis under ideal network conditions. The framework comprises six components: signal pre-processing and the CWRU dataset; robust non-IID data partitioning; the 1D-CNN local model; the aggregation method and federated learning framework; the communication cost model; and the multi-head DQN client selection agent. Fig. 1 presents the system design.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

3.1. Dataset Description and Signal Pre-processing

The experimental benchmark is the Case Western Reserve University (CWRU) bearing fault dataset. Four health conditions were used to gather vibration signals from a motor drive-end bearing: Normal — healthy baseline; Inner race fault; Outer race fault; and Ball fault. Signals were acquired at 12 kHz under four motor load conditions (0–3 HP) using fault sizes of 0.007, 0.014, and 0.021 inches. A sliding window of length L = 1,024 samples with no overlap is used to segment raw drive-end accelerometer signals. Every segment is separately normalized to have a unit variance and zero mean.

$$\hat{x} = \frac{(x - \mu)}{(\sigma + \varepsilon)} \tag{1}$$

where μ and σ are the per-segment mean and standard deviation, and ε is a small positive constant introduced to prevent division by zero. This per-segment normalization reduces amplitude variations across different load conditions. The resulting segments serve as direct input to the 1D-CNN with shape (1 × 1024).

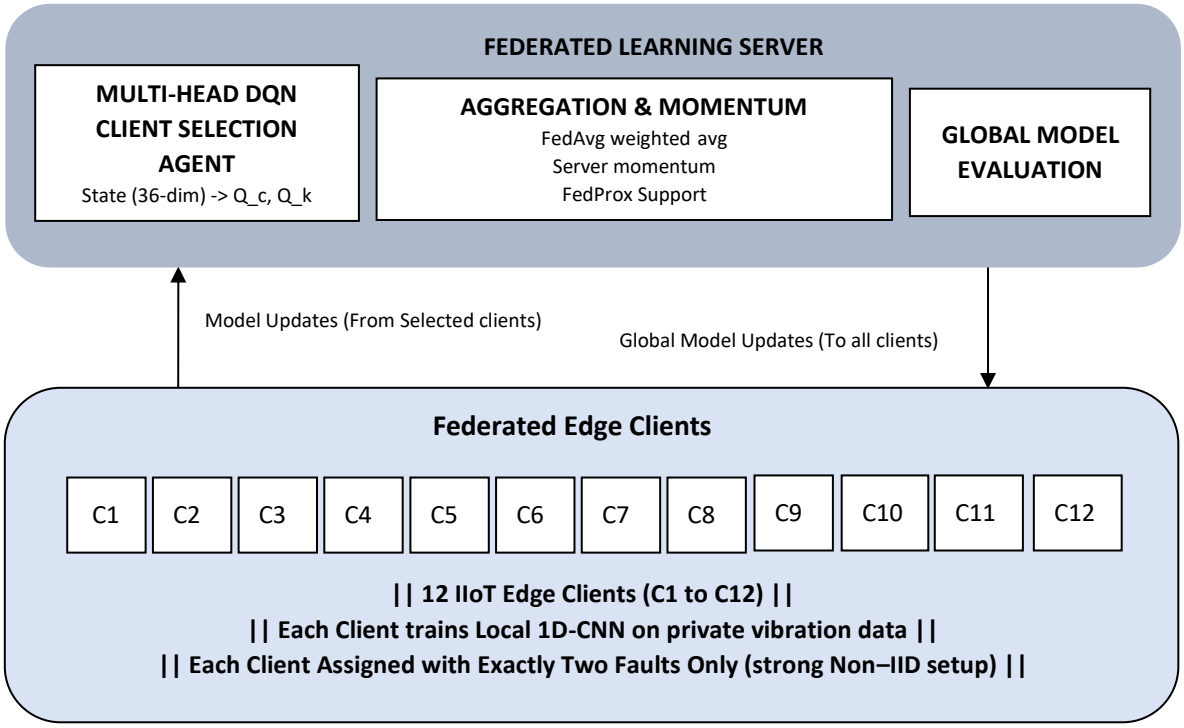

Fig. 1. Proposed Multi-Head DQN federated learning framework for bearing fault diagnosis.

3.2. Highly Non-IID Data Partitioning

Highly non-IID partitioning is used to accurately represent the data heterogeneity of actual IIoT systems. Data from precisely two of the four fault classes are allocated to each of the N = 12 federated clients. Since no client has samples from the other two classes, no single client can assess the global model across the entire distribution of fault classes. The 20% of segments from each vibration signal that belong to four faults under various fault sizes and load conditions are used to create the global server-side test set; no client is ever exposed to these segments. As a result, a balanced, multi-class test set including all four fault scenarios and all load states is generated, allowing an unbiased assessment of global accuracy using actual CWRU vibration signals.

3.3. Local Model: 1D Convolutional Neural Network

For local bearing fault classification, every federated client uses the same 1D-CNN architecture, which is intended to automatically extract discriminative temporal and spectral features from raw vibration segments without the need for manual feature engineering. Fig. 2 shows the architecture of the 1D-CNN. The Flower (flwr) federated learning framework is used for implementation, providing gRPC-based communication between the server and 12 client processes. Two local training objectives are evaluated: FedAvg and FedProx.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

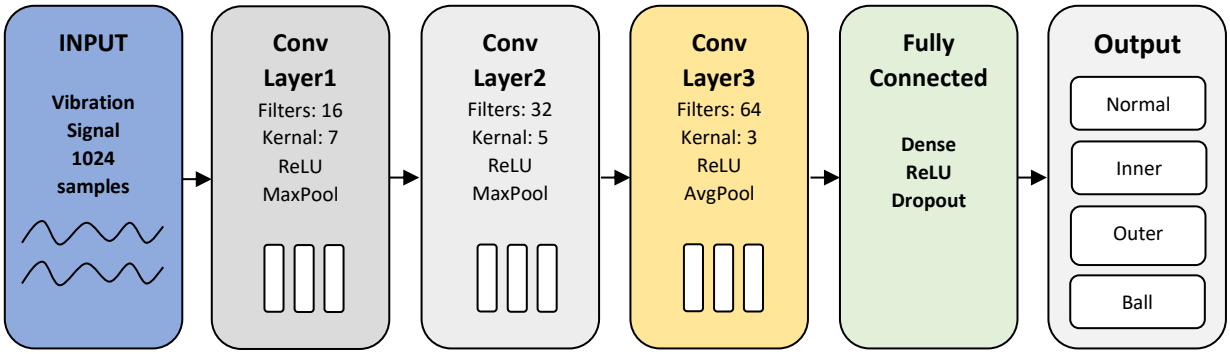

Fig. 2. One-dimensional CNN architecture for local model training.

FedAvg aggregation computes the weighted average of selected client updates:

$$\theta^{t+1} = \sum_{i \epsilon S_t} \left[\frac{n_i}{\sum_{j \epsilon S_t} n_j} \right] \theta_i^t \tag{2}$$

where S_t is the set of selected clients at communication round t, n_i is the number of local training samples of client i, and θ_i^t denotes the locally updated model parameters after E local training epochs, and θ^{t+1} is the aggregated global model. FedAvgM server momentum with β = 0.9 is applied to smooth aggregation under non-IID conditions.

$$\theta^{t+1} = \theta^t + \beta * \left(\theta_{agg}^{t+1} - \theta^t \right) \tag{3}$$

where θ_{agg}^{t+1} is the aggregated model obtained from Eq. (2), and β denotes the server momentum coefficient. FedProx local objective augments cross-entropy with a proximal term [8].

$$h_i(\theta) = F_i(\theta) + \left(\frac{\mu}{2} \right) * \left\| \theta - \theta_{global} \right\|^2, \mu = 0.5 \tag{4}$$

where $F_i(\theta)$ is the local empirical loss (cross-entropy) of client i, μ is the proximal coefficient, and θ_{global} denotes the current global model parameters. Because of the rigorous non-IID partitioning used in this work, the proximal term prevents local updates from deviating too far from the global model.

3.4. Communication Cost Model

A bidirectional communication cost consisting of a fixed download component and a variable upload component is incurred for every FL round.

$$C_t = (N + k_t) \times M \tag{5}$$

where N = 12 is the total number of clients (every round, every client receives the global model, resulting in a fixed download cost), $k_t \in \{6, 7, \ldots, 12\}$ is the number of chosen clients (variable upload cost), and M is the model size in MB. The communication savings attributable to reduced uploads, relative to the all-client baseline ($k_t = N$), are:

$$S_t^{\text{upload}} = \frac{(N - k_t)}{(N + k_t)} \tag{6}$$

At $k_t = 6$, the maximum communication saving per round is achieved, resulting in a maximum saving of 33.3% per round. Over T = 200 rounds, the total cost is given below.

$$C_{total} = \sum_{t=1}^{T} (N + k_t) \times M \tag{7}$$

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

3.5. Multi-Head Deep Q-Network for Client Selection

The Multi-Head DQN jointly determines the number of clients. $k_t \in \{6,\ldots,12\}$ and the identity of those clients at each FL round, treating client selection as a Markov decision process, and $N = 12$ is the total number of federated clients. Fig. 3 shows the architecture of the Multi-Head Deep Q-Network for joint client selection.

State space: At the beginning of round t, the server constructs a 36-dimensional state vector from three privacy-preserving statistics per client.

$$s_t = [\, acc_i, (1 - loss_i), n_i/1000 \,]_{i=1}^{N} \in \mathbb{R}^{36} \quad (8)$$

where acc_i and $loss_i$ are the local validation accuracy and loss of client i on the current global model, and n_i is the local training dataset size. All three features are observable by the server without accessing raw client data, preserving data privacy.

DQN architecture: The policy network consists of a shared feature extraction trunk and two specialized output heads. After processing the 36-dimensional state vector using the shared feature extraction trunk, the policy network (Q_policy) divides into two output heads: a client scoring head (Q_c) and a k-selection head (Q_k).

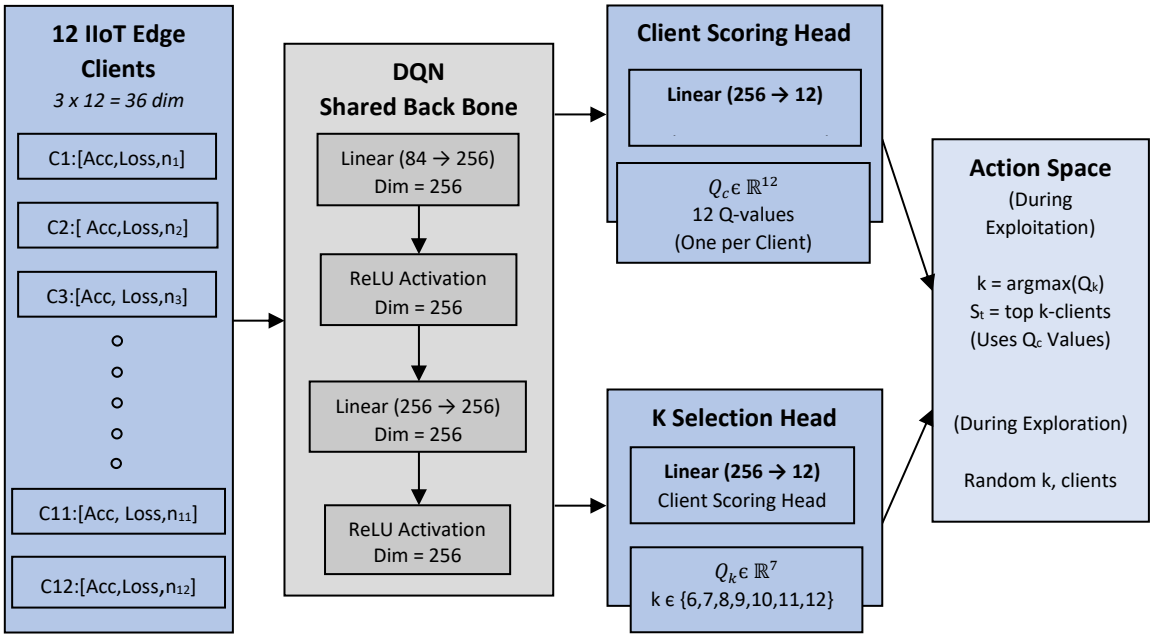

Fig. 3. Multi-Head DQN architecture for selecting k_t and the corresponding clients.

Reward function: The reward is designed to be mathematically consistent with the communication cost model and to optimize classification accuracy and communication efficiency simultaneously.

$$R_t = 0.50\, \Delta Acc_t \times 100 + 0.30\, S_t^{\text{upload}} \times 100 + 0.20\, Stab_t \times 100 \quad (9)$$

where $\Delta Acc_t = max(0, Acc_t - Acc_{t-1})$ is the round-over-round accuracy gain (50% weight — primary objective), $S_t^{\text{upload}} = (N - k_t)/(N + k_t)$ is the upload saving directly derived from Eq. (6) (30% weight); and $Stab_t = 1 - |S_t \triangle S_{t-1}|/N$ penalizes large round-to-round changes in the selected client set (20% weight).

4 Experimental Results and Discussion

4.1. Experimental Setup

All experiments are conducted on the CWRU bearing fault dataset with 4-class classification (Normal, Inner Race, Outer Race, Ball) across 12 federated clients under highly non-IID partitioning (each client holds exactly 2 classes). The Flower framework is used for all FL implementations with T = 200 communication rounds. Six FL configurations are compared: FedAvg, FedProx, FedAvg+FedRandom, FedProx+FedRandom, FedAvg+FedDQN, and FedProx+FedDQN (proposed). All methods use the same 1D-CNN architecture, Adam optimizer ($\eta = 0.001$), E = 15 local epochs, and B = 64 batch size.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

The evaluated configurations comprise the standard FedAvg [6] and FedProx [8] algorithms, together with four baseline and proposed variants that combine these aggregation methods with either random client selection or the proposed Multi-Head DQN-based client selection strategy. For DQN and Random methods, the client selection range is $k \in \{6, 7, \ldots, 12\}$. The experiments were conducted over five independent runs with different random seeds, and the average results are reported.

4.2. Global Test Accuracy Comparison

Fig. 4 compares the average global test accuracy of the six evaluated FL configurations at the final communication round (T = 200) over five independent runs. The proposed FedProx+FedDQN method achieves 94.05% ± 1.6% accuracy, outperforming FedAvg (88.46% ± 5.14%) and FedAvg+FedDQN (86.65% ± 4.78%), demonstrating the benefit of combining proximal optimization with intelligent client selection. FedProx + FedRandom achieves 95.33% ± 1.4%, marginally 1.28% higher than the proposed method, and this difference is smaller than the standard deviations of both methods; in addition, a paired t-test across five independent runs yields t = 1.722, p = 0.16 (> 0.05), confirming that the difference is not statistically significant. However, FedProx+Random reduces communication cost by 12.77%, whereas FedProx+FedDQN reduces it by 19.15%. This 6.38% additional communication saving over FedProx+FedRandom while maintaining almost equal accuracy demonstrates that DQN learns to optimize the multi-objective reward rather than maximizing accuracy alone. The proposed FedProx+FedDQN balances accuracy with stable and intelligent client selection, making it more reliable for practical deployment. Table 1 summarizes the accuracy of the six evaluated configurations across five runs.

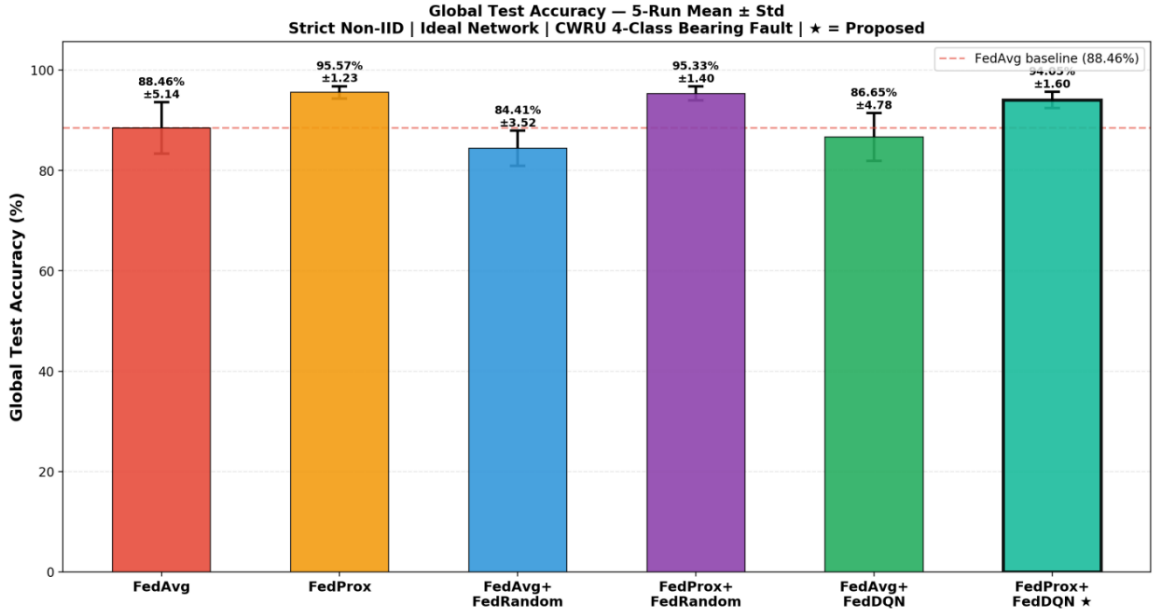

Fig. 4. Global test accuracy of the evaluated FL configurations.

Table 1. Global test accuracy of the evaluated FL configurations across five runs.

Method	Run1	Run2	Run3	Run4	Run5	Mean ± StdDev
FedAvg	79.63	91.52	88.22	91.63	91.3	88.46 ± 5.14
FedProx	94.82	97.03	93.94	95.7	96.37	95.57 ± 1.23
FedAvg+FedRandom	80.18	86.12	84.14	82.27	89.32	84.41 ± 3.52
FedProx+FedRandom	95.37	96.48	92.95	95.7	96.15	95.33 ± 1.40
FedAvg+FedDQN	89.21	89.98	88.55	87.22	78.3	86.65 ± 4.78
FedProx+FedDQN	92.51	96.7	93.72	93.17	94.16	94.05 ± 1.60

The proposed FedProx+FedDQN method achieves accuracy within a small margin of the FedProx baseline while selecting only. $k \in \{6, \ldots, 12\}$ clients per round, demonstrating that the Multi-Head DQN effectively identifies the most informative client subsets. Importantly, both DQN methods achieve performance comparable to their respective baseline methods (FedAvg and FedProx) while selecting fewer clients, confirming that learned client selection provides meaningful benefits under highly non-IID conditions. Table 2 presents the class-wise classification accuracy of FedProx with random client selection and the proposed FedProx+FedDQN method for one run.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

Compared with random client selection, the proposed DQN-based strategy improves the classification accuracy of the Outer Race fault and Ball fault classes while maintaining 100% accuracy for the Normal class. However, the Inner Race fault exhibits lower accuracy under the highly non-IID setting, indicating that the learned client selection policy emphasizes globally beneficial client updates rather than balanced representation of individual fault categories. This class-wise variation reflects the trade-off between communication efficiency and fault-specific recognition performance.

Table 2. Global test accuracy and per-class accuracy comparison (Final round, T=200).

Method	Acc (%)	Normal (%)	Inner (%)	Outer (%)	Ball (%)	Selection
FedAvg	88.22	100	77.78	48.61	99.31	All 12
FedProx	93.94	100	83.33	85.42	93.06	All 12
FedAvg+FedRandom	84.14	100	79.86	22.92	97.22	Random k
FedProx+FedRandom	92.95	100	81.25	77.78	96.53	Random k
FedAvg+FedDQN	88.55	100	68.75	59.03	100	DQN k
FedProx+FedDQN	93.72	100	61.81	99.31	99.31	DQN k

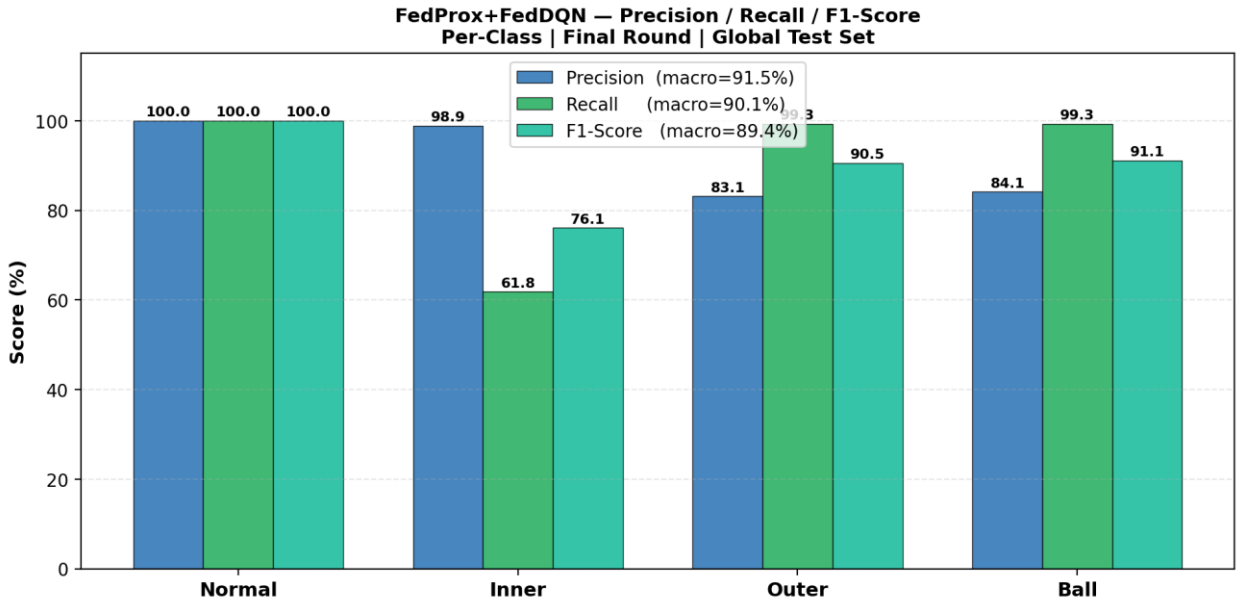

Fig. 5. Precision/Recall/F1-score of method FedProx+FedDQN.

Fig. 5 reports the precision, recall, and F1-score for the proposed FedProx+FedDQN at the final communication round (T = 200) for one experimental run. The communication performance of the proposed method is discussed in the next section.

4.3. Communication Cost Analysis

Table 3 presents the average cumulative communication cost (MB) over T = 200 rounds across five independent experimental runs for all federated learning methods, computed using the formula. $C_t = (N + k_t) \times M$ per round, where N = 12 (fixed download to all clients) and k_t is the number of selected clients (variable upload). The baseline is the all-client case: $C\_baseline = 2 \times N \times M \times T$.

Table 3. Cumulative communication cost over T=200 rounds.

Method	Total Comm (MB)	Baseline (MB)	Comm Saved (%)	Avg k / Round
FedAvg	243.19	-	-	12.00
FedProx	243.19	-	-	12.00
FedAvg+FedRandom	212.13	243.19	12.77	8.92
FedProx+FedRandom	212.13	243.19	12.77	8.92
FedAvg+FedDQN	195.35	243.19	19.67	7.26
FedProx+FedDQN	196.60	243.19	19.15	7.40

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

As shown in Table 3, FedAvg and FedProx incur the maximum communication cost since all 12 clients participate in every round. The DQN-based methods achieve significant savings by dynamically selecting fewer clients per round: FedProx+FedDQN reduces cumulative communication cost by 19.15% relative to the FedProx baseline, confirming the effectiveness of the Multi-Head DQN's joint optimization of k and client identity. Importantly, the DQN methods consistently outperform random selection (FedAvg+FedRandom, FedProx+FedRandom) in communication efficiency, demonstrating that the learned selection policy achieves a better trade-off between accuracy and communication cost than stochastic approaches.

It is noted that the communication savings arise exclusively from the upload component of the cost model: $S_t^{upload} = (N - k_t)/(N + k_t)$. At the minimum selection $k_t = 6$, the per-round saving is $6/18 \approx 33.3\%$. The DQN agent learns to balance this saving against accuracy maintenance, converging to an average k that maximizes the combined reward over 200 rounds. Table 4 reports the mean cumulative communication cost over 200 rounds with standard deviation across five independent runs.

Table 4. Cumulative cost across five independent experimental runs.

Method	Run1	Run2	Run3	Run4	Run5	Mean ± StdDev
FedAvg	243.19	243.19	243.19	243.19	243.19	243.19 ± 0
FedProx	243.19	243.19	243.19	243.19	243.19	243.19 ± 0
FedAvg+FedRandom	210.55	213.25	212.99	213.20	213.20	212.13 ± 1.4
FedProx+FedRandom	210.55	213.25	212.99	213.20	213.20	212.13 ± 1.4
FedAvg+FedDQN	188.96	194.81	193.29	203.98	203.98	195.35 ± 5.5
FedProx+FedDQN	188.96	194.8	193.29	197.56	208.41	196.60 ± 7.3

Table 4 shows that FedAvg and FedProx show zero standard deviation in communication cost since all 12 clients participate in every round, making the communication cost perfectly deterministic regardless of the random seed. Random selection methods show small variance ($\pm 1.4\ MB$) due to stochastic client count sampling across seeds. DQN-based methods exhibit higher variance ($\pm 7.3\ MB$) due to additional randomness in DQN training. However, DQN methods consistently achieve greater communication savings than random selection across all five seeds, confirming the robustness of the learned selection policy.

4.4. Convergence Analysis

Fig. 6 illustrates the convergence behavior of the six evaluated FL configurations over 200 FL rounds. Several key observations emerge from Fig. 6. First, FedProx converges more stably than FedAvg under strict non-IID conditions, consistent with its proximal regularization constraining local drift [8]. Second, random selection shows higher variance in accuracy across rounds compared to DQN selection, confirming that the Multi-Head DQN provides more stable aggregation by consistently selecting complementary client subsets.

4.5. DQN Client Selection Diagnostics

Fig. 7 illustrates the DQN client selection behavior over 200 rounds, including the value of k selected per round and the ε-decay schedule. The DQN agent exhibits three phases of client selection behavior. In the exploration phase (rounds 1–50), $\varepsilon \approx 1.0$ and k is sampled uniformly from $\{6, \ldots, 12\}$, yielding high variance. In the transition phase (rounds 50–150), ε decays and the agent begins to exploit learned Q-values, with k converging toward the range that maximizes the reward balance.

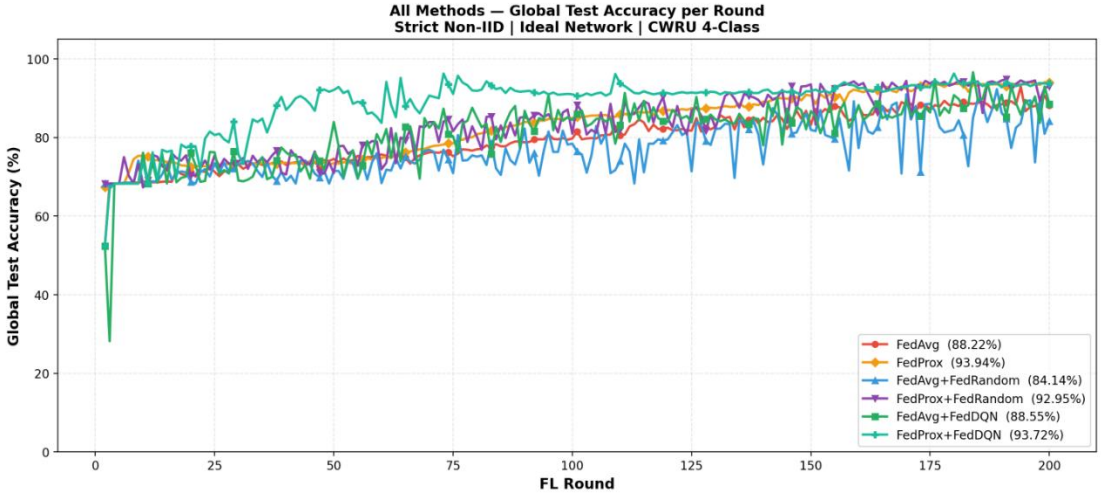

Fig. 6. Round-wise global test accuracy of the evaluated FL configurations.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

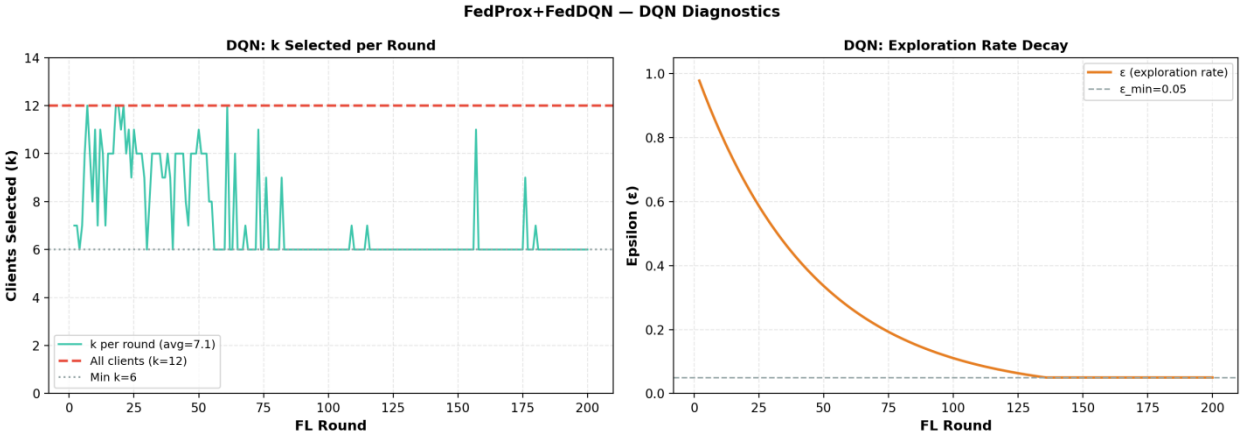

Fig. 7. Client selection using DQN

In the exploitation phase (rounds 150–200), $\varepsilon \approx 0.05$ and the agent consistently selects $k = 6$ clients per round — the range that provides the best accuracy/cost trade-off given the reward weights (0.50 × accuracy, 0.30 × upload saving, 0.20 × stability). The experimental results collectively demonstrate that the proposed Multi-Head DQN client selection framework achieves its primary objective: reducing communication cost while maintaining accuracy comparable to full-participation baselines.

4.6. DQN Overhead

The server-side inference and training times are measured to quantify the computational overhead introduced by executing the Multi-Head DQN at each FL round. The DQN inference—computing Q-values from the 36-dimensional state vector to select k and client identities—requires 0.0822 ms per round, and the 8 gradient update steps require 45.5461 ms per round, resulting in a total DQN overhead of 45.6283 ms per round. The estimated FL round computation time—comprising local training across 12 clients for 15 epochs—is approximately 1443.96 ms, making the DQN overhead only 3.16% of the total FL round time. The Multi-Head DQN model comprises 80,147 parameters (0.3057 MB), representing negligible server memory overhead compared to the 1D-CNN global model (0.58 MB). The proposed architecture is practically deployable in resource-constrained IIoT contexts without requiring any modification to current client hardware because the DQN operates entirely on the server and adds no extra computation or communication cost at the client side.

4.7. Comparison of Local Models: 1D-CNN and LSTM

A comparison between 1D-CNN and standard LSTM was performed as local models across all six FL methods while maintaining the same aggregation strategy, Multi-Head DQN client selection, reward function, and non-IID partitioning to assess the generalizability of the proposed framework across various local model architectures. Although the LSTM-based FedAvg model achieved the highest accuracy, it requires a significantly larger local model (5.56 MB). On the other hand, the 1D-CNN model in conjunction with FedProx reduces the model size to 0.58 MB and achieves similar diagnostic performance, resulting in more than a ninefold reduction in communication cost.

As a result, 1D-CNN with the proposed FedProx+FedDQN framework offers a better trade-off between accuracy and communication efficiency, making it better suited for IIoT installations with limited bandwidth. Table 5 presents the standard LSTM global model diagnostic accuracy and communication cost of all six methods over two experiments. As this is a representative study, two runs were considered sufficient to demonstrate architectural performance in terms of accuracy and communication cost. Since this paper's main contribution is the Multi-Head DQN client selection framework rather than the local model architecture, the five independent experimental runs are conducted only using 1D-CNN, which is the proposed and recommended local model.

Table 5. Performance of the six FL configurations using the standard LSTM local model.

	Run1			Run2		
Method	**Accuracy (%)**	**Communication Cost (MB)**	**Avg k**	**Accuracy (%)**	**Communication Cost (MB)**	**Avg k**
FedAvg	95.59	1163.1735	12.0	96.15	1163.1740	12.0
FedProx	93.61	1163.1735	12.0	87.00	1163.1740	12.0
FedAvg+FedRandom	97.36	1007.0608	8.8	94.16	1018.7500	8.9
FedProx+FedRandom	84.47	1007.0608	8.8	94.05	1018.7500	8.9
FedAvg+FedDQN	92.29	904.2846	6.7	89.54	925.4731	7.1
FedProx+FedDQN	93.72	903.7925	6.6	84.14	925.4731	**7.06**

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

4.8. Ablation Study

An ablation study was carried out over four weight configurations while maintaining the same 1D-CNN local model, Multi-Head DQN architecture, non-IID partitioning, and FL framework to support the design decisions of the proposed reward function. Configuration C2 (accuracy alone, 1.00/0.00/0.00) behaves similarly to full-participation FedAvg and negates the paper's key contribution of communication efficiency while achieving the highest accuracy but lower communication savings. Configuration C3 (Communication-only, 0.00/1.00/0.00) achieves accuracy comparable to the proposed method due to the minimum client constraint MIN_K = 6, which ensures fault class coverage under highly non-IID partitioning; however, relaxing this constraint would cause the DQN to aggressively select k < 6 clients, significantly degrading accuracy, confirming that the accuracy term (weight = 0.50) is an essential safeguard. Configuration C4 (stability-only, 0.00/0.00/1.00) decreases communication savings by pushing the DQN to choose all 12 clients each round, but it increases accuracy through consistent aggregation. The proposed weights (0.50/0.30/0.20) achieve a balanced trade-off: accuracy within 1–2% of the upper bound, 19.15% communication reduction, and stable non-IID convergence, confirming that all three reward components are necessary and their weights are carefully justified. Table 6 presents the ablation study results.

Table 6. Ablation study for reward components.

Configuration	Description	w_acc	w_save	w_stab	Final Acc (%)	Total Comm (MB)	Avg k
C1	Proposed (0.50/0.30/0.20)	0.5	0.3	0.2	93.7225	193.4905	7.1
C2	Acc Only (1.00/0.00/0.00)	1	0	0	95.3744	222.0559	9.91
C3	Comm Only (0.00/1.00/0.00)	0	1	0	93.6123	193.3378	7.08
C4	Stab Only (0.00/0.00/1.00)	0	0	1	96.0352	234.1236	11.11

To further validate the necessity of the accuracy term, an additional experiment with MIN_K = 1 demonstrates that Configuration C3 accuracy degrades to 68%. In comparison, Configuration C1 maintains 75%, confirming that the accuracy weight (0.50) serves as an essential safeguard independent of external constraints.

4.9. DQN Training Curves

The training curves for the FedProx+FedDQN approach across 200 communication rounds are shown in Fig. 8 to provide insight into the learning behavior of the proposed Multi-Head DQN agent. Fig. 8 comprises four panels illustrating different aspects of the DQN learning process

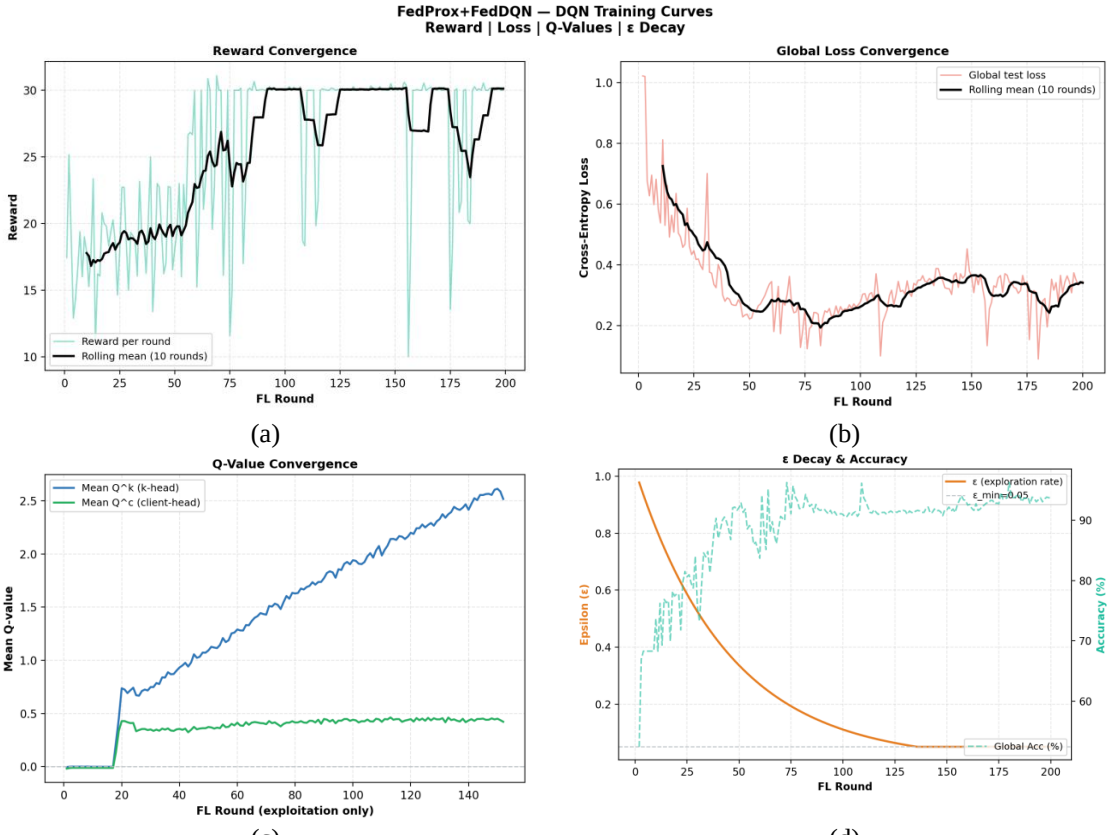

Fig. 8. DQN training curves

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

Fig. 8(a) presents the reward convergence with a 10-round rolling mean, illustrating the learning progress of the DQN agent. Fig. 8(b) shows the global cross-entropy loss, which steadily decreases as the global model improves. Fig. 8(c) presents the mean Q-value convergence for both the k-selection head (Q_k) and the client-scoring head (Q_c), indicating stable Q-value estimation after the exploration phase. Fig. 8(d) shows the epsilon decay schedule overlaid with the global accuracy, illustrating that accuracy improves as exploration decreases and exploitation increases. As shown in Fig. 8, the reward signal gradually rises from roughly 15 to 30 over 200 rounds, indicating that the DQN progressively learns a more effective client selection policy. In the first 75 rounds, the global cross-entropy loss rapidly drops and then stabilizes, indicating consistent model convergence under highly non-IID conditions.

The Q-value convergence demonstrates that while the client head. (Q^c) stabilizes early, the k-head (Q^k) continues to improve its client-count prediction. This suggests that determining the optimal number of participating clients is a more difficult learning problem than ranking client informativeness. The epsilon decay curve closely follows the improvement in global accuracy; as epsilon drops from 1.0 to 0.05, accuracy increases correspondingly. This validates the chosen decay schedule of 0.978 per round and confirms that the shift from exploration to exploitation is associated with a performance improvement. The present evaluation was conducted under ideal network conditions and on the CWRU benchmark dataset. Consequently, the reported performance does not account for network latency, packet loss, bandwidth variability, or additional industrial datasets, which remain important directions for future investigation.

5 Conclusion

This paper proposes a novel Multi-Head Deep Q-Network (Multi-Head DQN)-based client selection framework for communication-efficient federated learning applied to bearing fault diagnosis in IIoT environments. The framework addresses two fundamental limitations of existing FL approaches for IIoT fault diagnosis: the prohibitive communication overhead of full client participation in every round, and the inability of static or random selection strategies to adapt to the evolving global model state under highly non-IID data distributions. The proposed method was evaluated on the CWRU bearing fault dataset with 12 federated clients, each holding data from exactly two of the four fault classes—a highly non-IID partitioning that faithfully replicates the heterogeneity of real industrial deployments. The principal contribution of this work is a dual-head DQN architecture that jointly determines the optimal number of participating clients $k_t \in \{6,\ldots,12\}$ and their identities at each FL round, through a shared feature trunk feeding a client head (12 Q-values) and a k-head (7 Q-values). This joint optimization—not addressed by prior DQN-based FL selection methods—enables the agent to dynamically trade off communication cost against model quality in a single integrated decision.

A systematic six-method comparative study was conducted under the Flower FL framework, isolating the contributions of local training strategy (FedAvg vs FedProx), client selection mechanism (all clients, random, DQN), and their interaction under ideal network conditions. Experimental results on the CWRU benchmark dataset demonstrate that the proposed FedProx+FedDQN method achieves global fault classification accuracy comparable to the full-participation FedProx baseline, while reducing cumulative communication cost by 19.15% over 200 FL rounds. The method consistently outperforms random-selection counterparts in communication efficiency while achieving competitive classification performance, confirming that learned, adaptive client selection provides meaningful and reproducible gains over stochastic approaches under non-IID conditions.

The proposed FedProx+FedDQN combination further demonstrates that proximal regularization and intelligent selection are complementary: FedProx stabilizes convergence under the non-IID heterogeneity exacerbated by partial client participation, while the DQN ensures that the participating subset is maximally informative. This study was conducted under ideal network conditions (no packet loss, no latency, no bandwidth constraint). In real IIoT deployments, network impairments would introduce additional challenges, including gradient staleness and asynchronous client updates. Extending the Multi-Head DQN state space to include network quality indicators—as explored in network-aware FL frameworks—represents a natural direction for future work. Future work will also focus on incorporating class-aware reward functions into the DQN-based client selection strategy to improve the recognition of underrepresented fault classes while preserving communication efficiency.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ETHICS STATEMENT

This study did not involve human or animal subjects and, therefore, did not require ethical approval.

STATEMENT OF CONFLICT OF INTERESTS

The authors declare no conflicts of interest related to this study.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

REFERENCES

[1] Y. Yoo, H. Jo, and S.-W. Ban, "Lite and efficient deep learning model for bearing fault diagnosis using the CWRU dataset," *Sensors*, vol. 23, no. 6, p. 3157, Mar. 2023, doi: 10.3390/s23063157.

[2] K. Ma, S. Wu, D. Kong and X. Wang, "Rolling Bearing Fault Diagnosis Based on VMD-CNN-SLSTM Model," *2025 IEEE International Conference on Mechatronics and Automation (ICMA)*, Beijing, China, 2025, pp. 78-84, doi: 10.1109/ICMA65362.2025.11120800.

[3] X. Zhang, B. Qiu, S. Liu, M. Zhao, W. Li and X. V. Wang, "Intelligent Bearing Fault Diagnosis with Few-Shot Samples Based on Dynamics Simulation and A DataDriven Method," *2025 6th International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, Shanghai, China, 2025, pp. 491-495, doi: 10.1109/ICBAIE66852.2025.11326639.

[4] Z. Zhao *et al.*, "Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study," *ISA Transactions*, vol. 107, pp. 224–255, Aug. 2020, doi: 10.1016/j.isatra.2020.08.010.

[5] D. C. Nguyen *et al.*, "Federated learning for industrial internet of things in future industries," *IEEE Wireless Communications*, vol. 28, no. 6, pp. 192–199, Aug. 2021, doi: 10.1109/mwc.001.2100102.

[6] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proc. 20th Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, vol. 54, Apr. 2017, pp. 1273–1282.

[7] Z. Zhang, F. Zhou, C. Wang, C. Wen, X. Hu, and T. Wang, "A multiscale recursive attention gate federation method for multiple working conditions fault diagnosis," *Entropy*, vol. 25, no. 8, p. 1165, Aug. 2023, doi: 10.3390/e25081165.

[8] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated Optimization in Heterogeneous Networks," in *Proc. Machine Learning and Systems (MLSys)*, vol. 2, 2020, pp. 429–450.

[9] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li and H. Vincent Poor, "Federated Learning for Internet of Things: A Comprehensive Survey," in *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1622-1658, thirdquarter 2021, doi: 10.1109/COMST.2021.3075439.

[10] M. H. Mahmoud, A. Albaseer, M. Abdallah and N. Al-Dhahir, "Federated Learning Resource Optimization and Client Selection for Total Energy Minimization Under Outage, Latency, and Bandwidth Constraints With Partial or No CSI," in *IEEE Open Journal of the Communications Society*, vol. 4, pp. 936-953, 2023, doi: 10.1109/OJCOMS.2023.3263962.

[11] H. Wang, Z. Kaplan, D. Niu and B. Li, "Optimizing Federated Learning on Non-IID Data with Reinforcement Learning," *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, Toronto, ON, Canada, 2020, pp. 1698-1707, doi: 10.1109/INFOCOM41043.2020.9155494.

[12] Y. Xu, Z. Jiang, H. Xu, Z. Wang, C. Qian and C. Qiao, "Federated Learning With Client Selection and Gradient Compression in Heterogeneous Edge Systems," in *IEEE Transactions on Mobile Computing*, vol. 23, no. 5, pp. 5446-5461, May 2024, doi: 10.1109/TMC.2023.3309497.

[13] W. Sun, R. Yan, R. Jin, R. Zhao and Z. Chen, "FedAlign: Federated Model Alignment via Data-Free Knowledge Distillation for Machine Fault Diagnosis," in *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1-12, 2024, Art no. 3506112, doi: 10.1109/TIM.2023.3345910.

[14] F. Zhou, Y. Yang, C. Wang, and X. Hu, "Federated Learning Based fault diagnosis driven by Intra-Client Imbalance Degree," *Entropy*, vol. 25, no. 4, p. 606, Apr. 2023, doi: 10.3390/e25040606.

[15] J. Sun, F. Zhou, J. Chen, C. Wang, X. Hu, and T. Wang, "A federated adversarial fault diagnosis method driven by fault information discrepancy," *Entropy*, vol. 26, no. 9, p. 718, Aug. 2024, doi: 10.3390/e26090718.

[16] Z. Ke, Y. Yang, Z. Fan, Y. Zhou, H. Wang and Y. Liu, "Multi-Condition Rolling Bearing Fault Diagnosis Based on Prototype Contrast Federated Transfer Learning," *2025 Global Reliability and Prognostics and Health Management Conference (PHM-Xian)*, Xian, China, 2025, pp. 1-5, doi: 10.1109/PHM-Xian66756.2025.11427739.

[17] X. Ma, J. Zhu, Z. Lin, S. Chen, and Y. Qin, "A state-of-the-art survey on solving non-IID data in Federated Learning," *Future Generation Computer Systems*, vol. 135, pp. 244–258, May 2022, doi: 10.1016/j.future.2022.05.003.

[18] P. Kairouz *et al.*, "Advances and open problems in federated learning," *Foundations and Trends® in Machine Learning*, vol. 14, no. 1–2, pp. 1–210, Dec. 2020, doi: 10.1561/2200000083.

[19] S. P. Karimireddy, S. Kale, M. Mohri, S. U. Stich, A. T. Suresh, and S. Reddi, "SCAFFOLD: Stochastic Controlled Averaging for Federated Learning," in *Proc. 37th Int. Conf. Machine Learning (ICML)*, vol. 119, Nov. 2020, pp. 5132–5143.

[20] T. Nishio and R. Yonetani, "Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge," *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, Shanghai, China, 2019, pp. 1-7, doi: 10.1109/ICC.2019.8761315.

[21] Y. J. Cho, J. Wang, and G. Joshi, "Towards Understanding Biased Client Selection in Federated Learning," in *Proc. 25th Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, vol. 151, May 2022, pp. 10351–10375.

[22] Y. Jee Cho, S. Gupta, G. Joshi and O. Yağan, "Bandit-based Communication-Efficient Client Selection Strategies for Federated Learning," *2020 54th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, 2020, pp. 1066-1069, doi: 10.1109/IEEECONF51394.2020.9443523.

International Journal of Emerging Research in Engineering, Science, and Management
Vol. 5 | Issue 3 | pp. 67-80 | Jul-Sep 2026.
www.ijeresm.com | eISSN – 2583-4894
Received: 11 Apr 2026 | Revised: 17 Jul 2026 | Accepted: 26 Jul 2026 | Published: 07 Aug 2026

[23] X. Meng, Y. Li, J. Lu, and X. Ren, "An optimization method for Non-IID federated learning based on deep reinforcement learning," *Sensors*, vol. 23, no. 22, p. 9226, Nov. 2023, doi: 10.3390/s23229226.

[24] J. Wang, "FedCcs: Federated Learning Cluster-Based Client Selection Algorithm for Non-IID Data," *2025 5th International Conference on Neural Networks, Information and Communication Engineering (NNICE)*, Guangzhou, China, 2025, pp. 135-139, doi: 10.1109/NNICE64954.2025.11063913.