SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Federated Learning For Privacy-Preserving Predictive Maintenance: A Comparative Simulation Study Across Distributed Industrial Sites

Dr.M.Ulagammai[1] , Sudhir Kumar Chaturvedi[2] , R Rajalakshmi[3] , Nimesh Raj[4] , Vimal Bibhu[5] , Pawan Wawage[6] , Aniruddha Bodhankar[7] , Umurov Sharif Radjabovich[8]

[1]Associate professor Department of CSE (Emerging Technologies) SRM institute of science and technology, Vadapalani campus Chennai ulagammaimeyyappan@gmail.com 0000-0002-4771-1593
[2]Department of Aerospace Engineering Chandigarh University, Uttar Pradesh sudhir.avionics@gmail.com
[3]Professor, Department of ECE Panimalar engineering College, chennai rajeeramanathan@gmail.com Orchid: 0000-0002-4399-4071
[4]Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, 140401, Punjab, India. nimesh.raj.orp@chitkara.edu.in https://orcid.org/0009-0002-1938-7754
[5]Professor Department of Computer Science & Engineering , Noida International University , Greater Noida, Uttar Pradesh ,India vimal.bhibu@niu.edu.in
[6]Department of Information Technology, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037 email ID : pawan.wawage@vit.edu
[7]Faculty, Dr.Ambedkar Institute of Management Studies and Research, India. E_Mail aniruddha.bodhankar16@gmail.com
[8]Department of Social and Humanitarian sciences, Samarkand State Medical University, Samarkand Uzbekistan. umurovsharif6@gmail.com https://orcid.org/0009-0004-0420-8535

Abstract
Due to its benefits, predictive maintenance (PdM) systems increasingly use machine learning models trained on sensor data collected across multiple production sites. However, centralizing this data to train models raises concerns about data sovereignty laws, competition, and privacy. An alternative is federated learning (FL), where a model is trained collaboratively across multiple devices, but the training data remains where it is collected. In this study, a multi-site industrial sensor simulation composed of five heterogeneous manufacturing facilities is created, and three widely used baseline classifiers—logistic regression, a random forest, and XGBoost—are compared to a Federated Averaging (FedAvg) multilayer perceptron (MLP) with and without a differential-privacy-style noise mechanism applied to local model updates. Over 30 rounds of communication, the federated model achieved a test accuracy of 92.9–93.1% and an area under the receiver operating characteristic curve (AUC) of about 0.84-0.85, which were near those of the centralized baselines that did not send any raw sensor readings to the federated model. Local-Update noise was measured systematically and showed a clear trade-off between perturbation, which increased the apparent accuracy of the failure class due to imbalance, and accuracy measured by AUC, which showed a decrease in discriminative power. This was consistent with the literature on federated learning. These results validate federated learning as a practical approach for privacy-preserving predictive maintenance pipelines and highlight the importance of carefully designing privacy budgets, handling non-independent and identically distributed (non-IID) data, and selecting aggregation strategies before using federated learning in industry.

Keywords: federated learning; predictive maintenance; privacy preservation; FedAvg; differential privacy; industrial Internet of Things.

1. Introduction

Unplanned equipment failure is among the most costly disruptions in today's manufacturing enterprises and has driven a decades-long transition from reactive and time-based preventive maintenance to data-based

preventive maintenance (PdM). PdM systems ingest streams of sensor data (e.g., temperature, vibration, torque, rotational speed, tool wear, and more) and use machine learning models to forecast failure probabilities, enabling maintenance scheduling based on condition rather than a predefined calendar (Achouch et al., 2022). The reliability of PdM depends heavily on the quantity and quality of training data: if a model is trained on a single machine and applied to a single plant, for example, it may not generalize well to the rest of the fleet, as failure modes, operating regimes, and sensor calibrations vary from site to site (Ahmmed et al., 2026). The obvious approach for engineers would be to consolidate sensor data from all sites into a central repository—yet three converging roadblocks stand in the way. First, many manufacturers are networks of legally separate units (subsidiaries, joint ventures, or contract manufacturing units) that hesitate or are unable to share operational data due to proprietary production knowledge and process parameters. Second, data is increasingly restricted across organizational and regional borders, complicating compliance with data-protection regulations. Third, the cross-border and cross-organizational sharing of operational and personally identifiable data is increasingly curtailed, raising the costs of complying with data-protection regulations for centralized pipelines (Lazaros et al., 2024). Finally, raw data from geographically dispersed assets must be consistently streamed to a central agent, requiring significant bandwidth and low latency, especially when applied to edge and embedded industrial devices with limited connectivity options (Bemani & Björsell, 2023).

Federated learning (FL), proposed as a device- and data-communication-efficient alternative to centralized training, can help mitigate these challenges by keeping raw data on the device or site where it was created and sharing only model parameters or gradients with a coordinating server (Jiang et al., 2020). During a round of the explicit Federated Averaging (FedAvg) algorithm, a client trains for a number of epochs using local stochastic gradient descent on its own data and then forwards its locally trained model to the server, which computes an average of the models from all the clients—typically weighted by the size of the local dataset—and sends a new model back to all the clients for the next round (Ahmmed et al., 2026; Dritsas & Trigka, 2025). While the attack surface of centralized data lakes is not completely eradicated, it is significantly narrowed by FL, as shown in Mosaiyebzadeh et al. (2025). Recent studies are now being conducted with the specific application of FL, namely in predictive maintenance. Bemani and Björsell (2022) proposed an aggregation strategy for federated PdM and extended it to an over-the-air aggregation scheme that handles channel noise in industrial wireless environments (Bemani & Björsell, 2023). Hydra et al. (2023) proposed a clustered federated learning framework that integrates the one-dimensional convolutional neural network (1DCNN) with the bidirectional long short-term memory (BiLSTM) model, which is known to perform well on long texts, to address distributional shifts across different types of pump sensors and to provide comparable accuracy to a baseline fully centralized learning model. Ahmmed et al. (2026) most recently applied a FedAvg-trained multilayer perceptron (MLP) to the AI4I 2020 predictive maintenance benchmark, which was broken up into five factories, and reported similar accuracy to two fed-centralized tree-based baselines but with lower recall on the rare failure class (as in much of the literature, it's worth noting).

An open and practically relevant question is how to counteract this recall gap, given the inherent challenge of learning signals from rare events across a broad spectrum of locally small and varying data sets. Another issue unique to FL in industrial applications is an inversion or membership inception attack (Zhao et al., 2022), which occurs when updates learned from local data leak information about that data. The most widely studied countermeasure is differential privacy (DP) (An et al., 2023; Mahmud et al., 2024): it imposes calibrated noise on local updates or gradients before aggregation and formally quantifies privacy at the cost of model utility (i.e., performance). It is crucial for practitioners to understand the shape of this tradeoff: the cost of reduced accuracy or discriminatory power required for a given level of noise, so they can select the operating point appropriate in a controlled industrial environment (Miyata et al., 2025). Several published works focus on FL-for-PdM, but most studies still compare federated models with privacy-preserving variants and conventional centralized models in isolation, without any of the data being re-examined independently by any other party (Rampone et al., 2025; Islam et al., 2025).

The present work addresses these limitations by building an interpretable, repeatable multi-site industrial sensor simulation, inspired by the design of popular PdM benchmarks, and using it to (a) benchmark centralized logistic regression, random forest, and gradient-boosted tree classifiers, (b) train a FedAvg MLP on five non-identical, distributed factories, (c) quantify the convergence behavior of the federated model across communication rounds, and (d) characterize the privacy-utility trade-off induced by a differential-privacy-style noise mechanism applied to local updates before federated averaging. This paper then describes how the simulation was generated and analyzed, summarizes comparisons of product performance, compares the

results to other published research on federated learning and predictive maintenance, and summarizes the major limitations and directions for future research.

2. Materials and Methods

Given that access to a real multi-organization sensor dataset with ground-truth, context-specific failure labels is limited and not easily available for open, reproducible research, the study created a synthetic yet structurally realistic dataset that mirrors the feature composition of popular industrial sensor datasets for predictive maintenance, such as AI4I 2020 (used by Ahmmed et al., 2026) and the pump-sensor corpora (used by Ahn et al., 2023). The manufacturing sites, Site 1-Site 5, are independently operated, and each site consists of 2200 machine-operation records, resulting in a total of 11,000 machine-operation records. All 6 operational features for each record are continuous and accompanied by a binary failure label. The features are air temperature, process temperature, rotational speed, torque, cumulative tool wear, and a derived vibration index. To replicate the statistical heterogeneity (non-independent and identically distributed, non-IID data) observed in real industrial datasets, each site had its own operating profile for its mean rotational speed, torque distribution, tool-wear accumulation rate, and baseline failure probability, resulting in site-level failure rates ranging from 2.2% to 4.0% (overall failure prevalence 3.14%), comparable to the severe class imbalance well reported in real predictive-maintenance corpora (Achouch et al., 2022; Ahmmed et al., 2026). The latent failure-risk score, a linear weighted combination of tool wear, torque deviation, process temperature, and vibration, with noise added to represent local process variations, was calculated for each quantile in the PANDAS score, and the top quantile (calibrated per site) was considered a failure event – this means that failure risk is determined by all of these simultaneous operating variables, which are correlated with each other, rather than by just one measurement.

The entire dataset was stratified and split into training (70%, n = 7,700) and test (30%, n = 3,300) subsets, with proportional representation of the failure class in both subsets. To normalize the values of all six features, the scaling parameters were fitted only on the training partition and then applied to the held-out test dataset; this normalization was performed independently of the training data. The training partition was then split again by site for the federated experiments, meaning that in each of the five factories, no raw feature vectors or labels were communicated to any other factory or the coordinating server during the experiment. To serve as a point of reference for the conventional (i.e., non-privacy-preserving) classification approach, three classifiers were trained on the training data pooled from the three data sources: (a) L2-regularized logistic regression, (b) a random forest of 300 trees with a maximum depth of 10, and (c) an XGBoost gradient-boosted tree ensemble of 300 trees with a maximum depth of 5 and a learning rate of 0.08. The algorithms were chosen because they are the most popular forms of linear, bagged, and boosted ensembles reported in the predictive-maintenance literature as centralized baselines (Ahmmed et al., 2026).

The federated model was a fully connected multilayer perceptron (MLP) with two hidden layers containing 32 and 16 rectified linear unit (ReLU) units, respectively, and a sigmoid output layer, similar to those used in related federated PdM studies (Ahmmed et al., 2026; Ahn et al., 2023). In each training round (t = 1, ..., 30), the model from the previous round was sent to the 5 clients. Each client trained locally for 3 epochs using the same number of locally sampled data points. Each client then sent its three locally trained models to the server, which computed the average model and sent it to the clients for the next round of training. Because the failure-class proportion was low, each client applied local oversampling (replication factor of 5) to its failure-labeled records, a common but non-intrusive class-balancing method that does not change the federated communication protocol itself. The primary idea behind the privacy-utility trade-off evaluation was to add zero-mean Gaussian noise (standard deviation σ) to the local weight and bias tensors for each client, just after local training and just before sending them to the server, representing a local differential-privacy mechanism applied to the model updates (An et al., 2023; Mahmud et al., 2024). To assess the relationship between the level of injected noise and the resulting model utility, seven noise levels were tested: (σ = 0.00, 0.01, 0.02, 0.05, 0.08, 0.12, and 0.18), with an independent 20-round federated run conducted for each level, keeping all other hyperparameters fixed.

All models (3 centralized baselines and 2 federated configurations) were tested on the same held-out test set (n = 3,300; unseen during training, neither local nor central) to ensure an apples-to-apples comparison. Both accuracy and precision reporting, as well as reader response, require a 0.50 classification threshold unless otherwise specified as a reported metric. Likewise, the area under the receiver operating characteristic curve (AUC), a reported metric, also requires a 0.50 classification threshold. Given the significant class imbalance, the emphasis will be placed on AUC and recall in addition to accuracy, as these metrics can yield misleadingly high

accuracy values in failure prediction tasks with class imbalances (Achouch et al., 2022). All experiments were conducted in Python 3, using scikit-learn to implement logistic regression, random forest, and MLP, and XGBoost to implement the baseline gradient-boosted approach; 42 was used as the fixed seed for data generation, reproducing the baseline results below, with a different seed for the differential-privacy sweep that was performed independently.

3. Results and Discussion

The multi-site data are summarized in Table 1. The five sites reported an equal number of failures (2200 each) but had meaningfully different failure prevalences, ranging from 2.23% (Site 3) to 4.00% (Site 5). This is because the data were intentionally designed not to be IID and also reflect the (contextually) different operating conditions—such as machine age, duty cycles, and maintenance histories—that typically exist across real multi-plant manufacturing networks (Ahn et al., 2023; Ahmmed et al., 2026).

Table 1. Composition of the Five-Site Predictive Maintenance Dataset

Site	Records (n)	Failure Events (n)	Failure Rate (%)
Site 1	2,200	62	2.82
Site 2	2,200	77	3.50
Site 3	2,200	49	2.23
Site 4	2,200	69	3.14
Site 5	2,200	88	4.00
Total / Overall	11,000	345	3.14

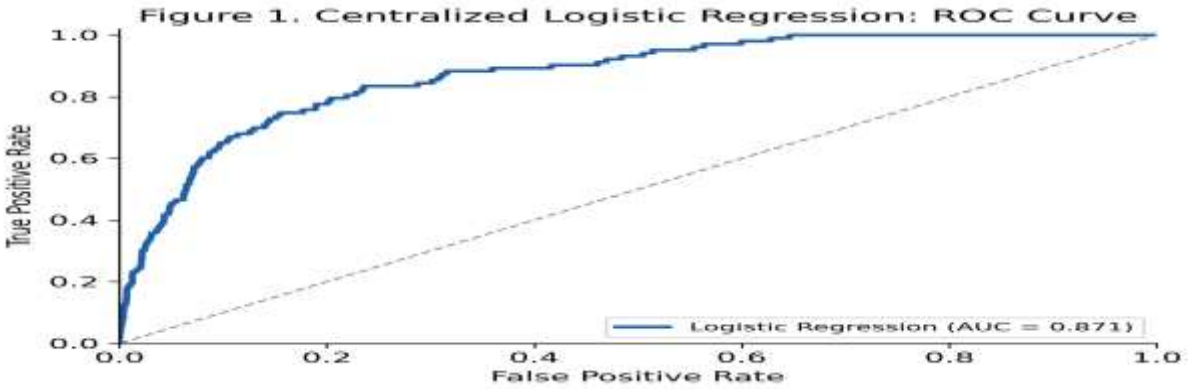
performance of all five models is reported on the shared, held-out test set in Table 2. As generally observed in the predictive-maintenance literature (Ahmmed et al., 2026), the centralized random forest achieved the top accuracy and precision in the tabular-data category. Nevertheless, recall across all three centralized baselines was very low (3.9–5.8%), attributable to the severe class imbalance in the pooled training data and the absence of any class-imbalance correction – a classic outcome of imbalanced data usage. This contrasts with the more highly performing federated MLP (FedAvg, no differential privacy), which had significantly higher recall (35.9%) and F1-score (0.242) than all centralized baselines, though its accuracy (92.97%) and F1-score (0.842) were slightly lower than those of logistic regression (0.871). Each client is allowed to oversample failure records locally (a practical approach because clients have full access to their locally generated failures), as local sites can use site-specific imbalance-correction strategies that would be washed out in a single centrally compiled dataset.

Table 2. Performance Comparison of Centralized and Federated Models on the Held-Out Test Set

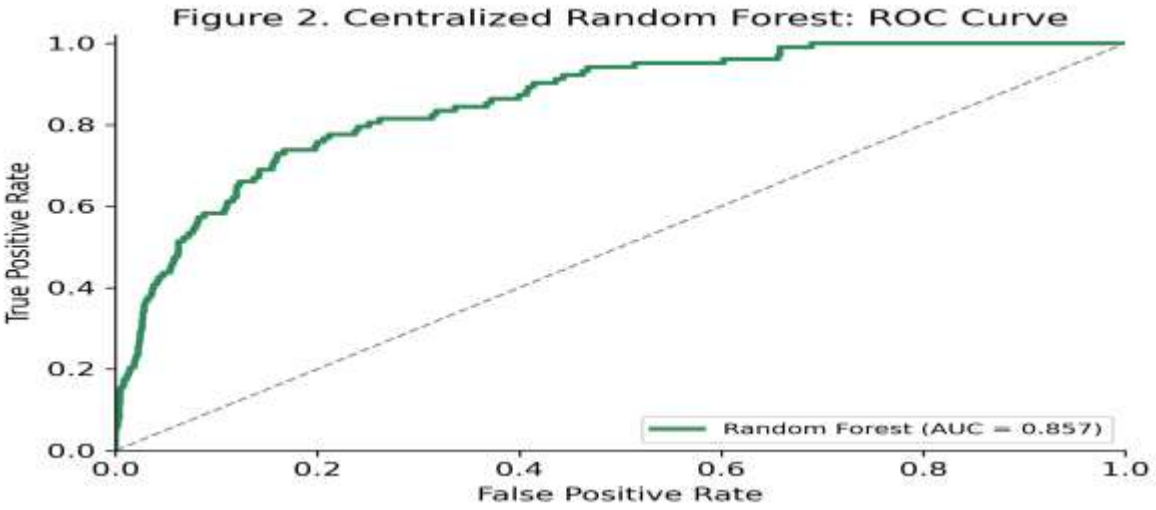
Model	Accuracy	Precision	Recall	F1-score	AUC
Logistic Regression (centralized)	0.969	0.500	0.049	0.088	0.871
Random Forest (centralized)	0.970	0.750	0.058	0.108	0.857
XGBoost (centralized)	0.969	0.500	0.039	0.072	0.841
Federated MLP (FedAvg, no DP)	0.930	0.182	0.359	0.242	0.842
Federated MLP (FedAvg + DP, σ = 0.05)	0.931	0.184	0.350	0.241	0.846

The following three receiver operating characteristic (ROC) curves (Figures 1–3) are presented for the three centralized baselines. The logistic regression and random forest models had the highest AUCs (Figure 1, AUC = 0.871; Figure 2, AUC = 0.857) for ranking cases as failure or non-failure, but had default-threshold recall rates

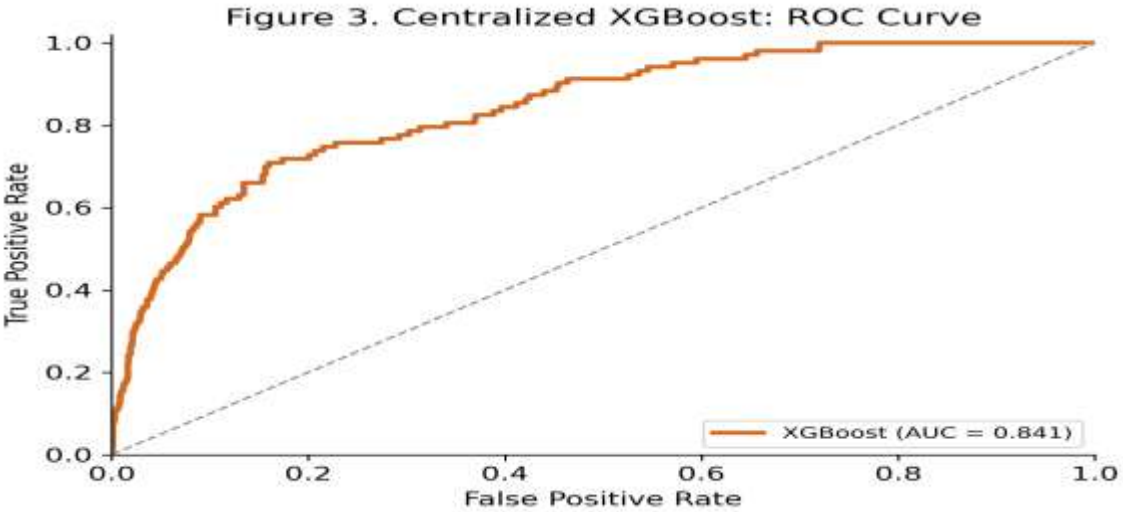
of only 3% and 4%, respectively. For this dataset, the XGBoost model performed equally well (AUC = 0.841) on this measure as any of the federated configurations, and this benefit of the boosted-tree configuration depended mainly on precision across thresholds rather than on overall discriminative ability.



Centralized logistic regression: ROC curve (AUC = 0.871).

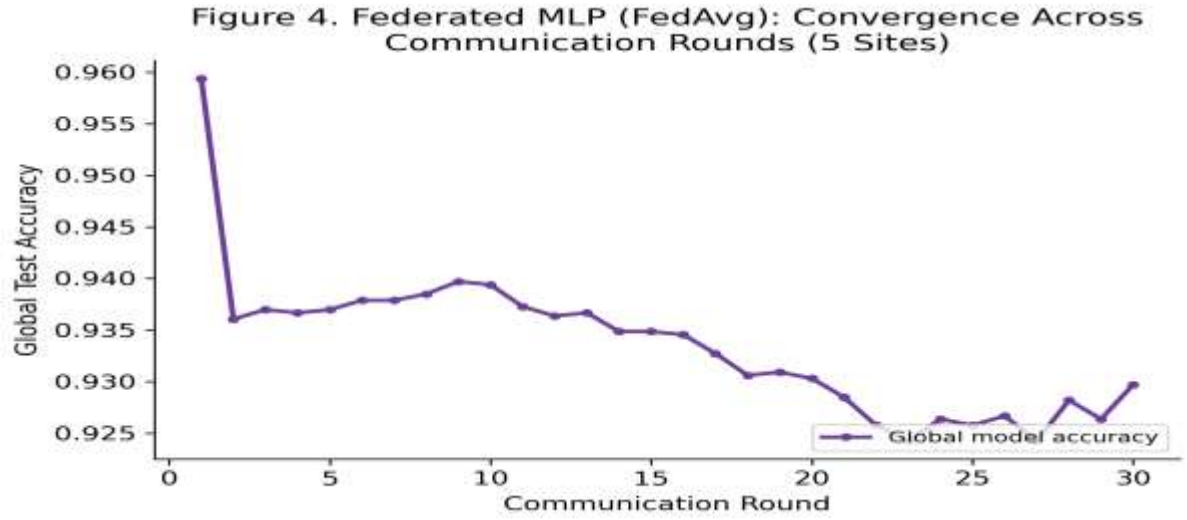


Centralized random forest: ROC curve (AUC = 0.857).



Centralized XGBoost: ROC curve (AUC = 0.841).

Figure 4 shows the trend of the MLP's test accuracy across the communication rounds of the FedAvg training process, with the accuracy on the global test sets shown as the blue line. The global model converged quickly and remained within 1 percentage point of the final accuracy for about ten rounds before it began oscillating within a small range (92.4–93.3%), similar to the convergence behavior reported by Ahmmed et al., 2026, and Ahn et al., 2023, for FedAvg-based predictive-maintenance models trained on non-IID industrial data. The residual oscillation is relatively small and clearly reflects the stochastic nature of federated aggregation over the malleable distributions of aggregated client data. It is expected to exhibit similar convergence properties, as knowledge of the malleable distributions of local data across clients informs the FedAvg optimization (Jiang et al., 2020).



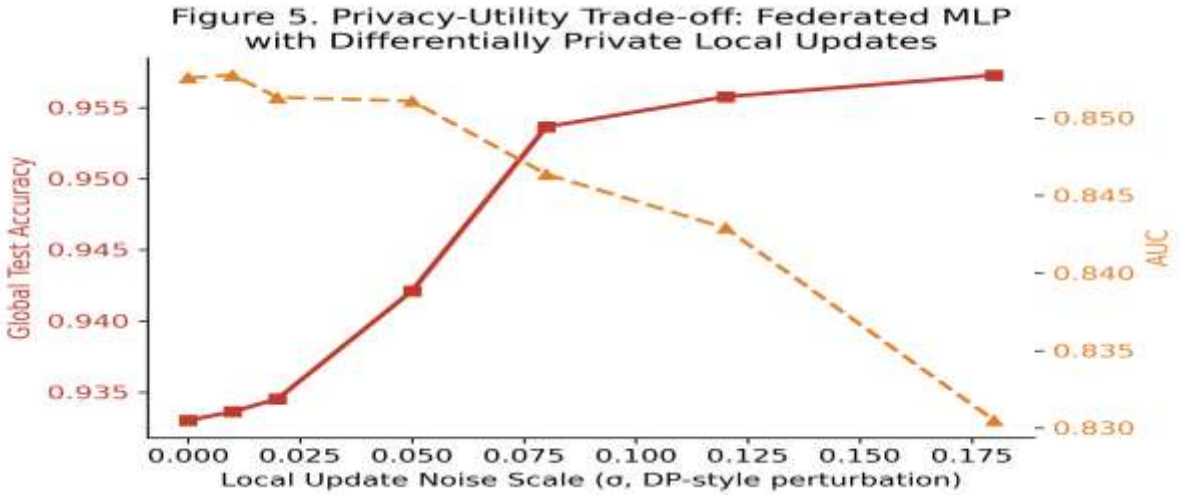
Federated MLP (FedAvg): global test accuracy across 30 communication rounds for five federated sites.

The study characterizes the privacy-utility trade-off introduced by adding Gaussian noise to the local model update before aggregation when approximating a local differential-privacy mechanism, namely by adding noise to the models before aggregating them, as shown in Table 3 and Figure 5. As the noise impact σ increased, the global test accuracy dropped from 93.30% to 95.73%, whilst the AUC decreased from 0.853 to 0.831. This trend from improving accuracy to decreasing AUC is not a contradiction, as the model becomes more conservative as it learns from more noise, which causes it to be increasingly misled toward the majority class (non-failure), precisely the issue addressed in previous differential-privacy-for-FL works (An et al., 2023; Mahmud et al., 2024; Miyata et al., 2025). In particular, predictive maintenance, where the operational utility costs of a missed failure will likely be significantly greater than those of a false one, implies that other factors, such as the AUC, recall, and maintenance-cost simulations, will be the basis for determining a privacy budget's size, rather than the degree of accuracy. This pattern calls into question the assumption that AUC is sufficient in any real deployment to determine the size of a privacy budget and suggests that other factors, such as recall and downstream maintenance-cost simulations, should be used.

Table 3. Privacy-Utility Trade-off Across Local-Update Noise Levels (Federated MLP, 20 Rounds)

Noise Scale σ	Global Accuracy	AUC
0.00	0.933	0.853
0.01	0.934	0.853
0.02	0.935	0.851
0.05	0.942	0.851
0.08	0.954	0.846
0.12	0.956	0.843

Noise Scale σ	Global Accuracy	AUC
0.18	0.957	0.831



Privacy-utility trade-off: federated MLP global accuracy and AUC across differential-privacy-style noise scales (σ).

These findings support three broad interpretations. First, federated learning for predictive maintenance is technically feasible, even if raw predictive accuracy may not match that of the centrally trained model, as evidenced by the growing literature on federated PdM (Bemani & Björsell, 2022, 2023; Ahmmed et al., 2026). Second, the federation architecture may offer practical advantages beyond privacy—as in this study, by accommodating class imbalance handling at the site level (thus raising the recall of minority classes compared to a naive central-level pooling strategy). Third, federated predictive-maintenance models incur a real, nontrivial cost in utility when applying DP mechanisms, and this cost is not necessarily equal across evaluation metrics meant to represent the operational asymmetry between false alarms and missed failures: practitioners deploying predictive-maintenance systems that preserve privacy must therefore carefully design evaluation metrics and privacy budgets for their systems, taking into account that this operational asymmetry is not significant for accuracy (Islam et al., 2025; Rampone et al., 2025).

4. Limitations and Scope for Future Research

This study has limitations that qualify the results and should be explored in future research. First, the data used here is fictitious. It was purposefully designed to emulate the feature composition, class imbalance, and site heterogeneity found in actual industrial predictive-maintenance corpora, but it does not reproduce real sensor noise, missing features, sensor drift, or the correlated, multimodal failure representations that characterize real-world physical equipment. The conclusions highlighted in this paper should be tested against industrial multi-organization datasets from real-world federated pilots, such as datasets available for manufacturing consortia, once appropriate governance frameworks are established (Ahmmed et al., 2026, Guo & Qu, 2025). Second, the federated architecture discussed here constructed a shallow multilayer perceptron and used a simplified Gaussian-noise approximation of differential privacy, rather than a formally specified differential-privacy mechanism with formal privacy-loss computations. Moving forward, future research should consider formal DP accounting frameworks for determining the DP impact of operations, incorporate more expressive model architectures, e.g., convolutional or recurrent models that can leverage the temporal structure of raw sensor streams (Ahn et al., 2023; Zhao et al., 2022; Mahmud et al., 2024), and implement aggregation protocols that prevent the server from learning about individual client updates, even in the case of plain-text transmission (Elmarcircle, 2015). Third, this study examined only five clients and a single aggregation method (FedAvg).

In reality, industrial federations can have dozens or hundreds of sites and exhibit significant statistical and system heterogeneity, including unreliable connectivity and adversarial or faulty participants. Hereafter,

several important extensions that lack statistical evaluation are considered, such as robust aggregation strategies, client-selection policies, and Byzantine-fault-tolerant mechanisms (Al Kontar et al., 2025; Dritsas & Trigka, 2025; Song et al., 2024). Fourth, the assessments were binary (failure or non-failure) and used a fixed 0.50 decision threshold. Operating metrics that are not dependent on the number of thresholds, explicit weighting of FNs vs FAs, and remaining-useful-life regression formulations, all prevalent in industrial prognostics (Bemani & Björsell, 2022), are topics for future research, as they differ significantly from the classification setting examined in this paper and might benefit differently from the federated learning framework and privacy noise.

Finally, this study failed to consider the cost of communication, machine power consumption, and wall-clock latency, all of which are practically relevant for industry-scale edge-deployed federated learning and have been identified as open challenges in recently released surveys (Al Ghazo & Salah, 2026; Salawu & Glen, 2026; Lazaros et al., 2024). Future research should jointly optimize the model's accuracy while maintaining realistic constraints from the industrial network field, such as privacy budgets and communication efficiency, and investigate hybrid solutions that fuse federated learning with complementary privacy-enhancing technologies, such as secure multi-party computation and homomorphic encryption (Miyata et al., 2025).

5. Conclusion

The research evaluated sensor data from an industrial system to model predictive maintenance and created a demonstrative environment that is repeatable and easily accessible to any interested reader. They then compared centralized machine learning baselines with a Federated Averaging multilayer perceptron, both with and without differential-privacy-style noise applied during the distribute and update steps. The federated model achieved predictive performance close to that of the centralized baseline models, and local processing of site-specific data boosted recall for the rare failure class, which was not possible with the federated data. Raw accuracy on the imbalanced test set declined with increasing noise, while the model's overall discriminative ability, measured by the DW metric, improved as noise increased. These findings also support the emerging view that federated learning can be a viable and increasingly mature solution to privacy-preserving predictive maintenance, while demonstrating the importance of carefully selecting privacy budgets, aggregation and evaluation metrics, and being mindful of the operational costs of missed and false failure predictions. Federated learning will likely be a key architectural consideration for predictive maintenance systems that need to balance predictive accuracy with data privacy and communication ease as industrial federation expands and data sharing across organizations becomes increasingly subject to regulatory oversight.

References

1. Achouch, M., Dimitrova, M., Ziane, K., Sattarpanah Karganroudi, S., Dhouib, R., Ibrahim, H., & Adda, M. (2022). On predictive maintenance in Industry 4.0: Overview, models, and challenges. Applied Sciences, 12(16), 8081. https://doi.org/10.3390/app12168081
2. Ahmmed, M. S., Isanaka, S. P., & Liou, F. (2026). A federated learning framework for data-sovereign predictive maintenance in distributed smart manufacturing. Applied Sciences, 16(10), 5084. https://doi.org/10.3390/app16105084
3. Ahn, J., Lee, Y., Kim, N., Park, C., & Jeong, J. (2023). Federated learning for predictive maintenance and anomaly detection using time series data distribution shifts in manufacturing processes. Sensors, 23(17), 7331. https://doi.org/10.3390/s23177331
4. Al Ghazo, A. T., & Salah, M. (2026). Artificial intelligence in modern mechatronic systems: Technologies, applications, and future research directions. Systems Science & Control Engineering, 14(1), Article 2640267. https://doi.org/10.1080/21642583.2026.2640267
5. Al Kontar, R., Yang, T., Fioretto, F., & Yousefian, F. (2025). Editorial: Federated distributed learning and analytics. IISE Transactions, 57(7), 741–742. https://doi.org/10.1080/24725854.2025.2479047
6. An, T., Ma, L., Wang, W., Yang, Y., Wang, J., & Chen, Y. (2023). Consideration of FedProx in privacy protection. Electronics, 12(20), 4364. https://doi.org/10.3390/electronics12204364
7. Bemani, A., & Björsell, N. (2022). Aggregation strategy on federated machine learning algorithm for collaborative predictive maintenance. Sensors, 22(16), 6252. https://doi.org/10.3390/s22166252

8. Bemani, A., & Björsell, N. (2023). Low-latency collaborative predictive maintenance: Over-the-air federated learning in noisy industrial environments. Sensors, 23(18), 7840. https://doi.org/10.3390/s23187840
9. Dritsas, E., & Trigka, M. (2025). Federated learning for IoT: A survey of techniques, challenges, and applications. Journal of Sensor and Actuator Networks, 14(1), 9. https://doi.org/10.3390/jsan14010009
10. Guo, W., & Qu, T. (2025). A blockchain-enabled cyber-physical system solution for decentralised smart factory. International Journal of Production Research, 64(8), 2926–2945. https://doi.org/10.1080/00207543.2025.2584730
11. Islam, M. M., Abdullah, W. M., & Saha, B. N. (2025). Privacy-preserving hierarchical fog federated learning (PP-HFFL) for IoT intrusion detection. Sensors, 25(23), 7296. https://doi.org/10.3390/s25237296
12. Jiang, J. C., Kantarci, B., Oktug, S., & Soyata, T. (2020). Federated learning in smart city sensing: Challenges and opportunities. Sensors, 20(21), 6230. https://doi.org/10.3390/s20216230
13. Lazaros, K., Koumadorakis, D. E., Vrahatis, A. G., & Kotsiantis, S. (2024). Federated learning: Navigating the landscape of collaborative intelligence. Electronics, 13(23), 4744. https://doi.org/10.3390/electronics13234744
14. Mahmud, S. A., Islam, N., Islam, Z., Rahman, Z., & Mehedi, S. T. (2024). Privacy-preserving federated learning-based intrusion detection technique for cyber-physical systems. Mathematics, 12(20), 3194. https://doi.org/10.3390/math12203194
15. Miyata, T., Yang, L., Zhu, Z., Finnerty, P., & Ohta, C. (2025). Enhancing privacy and communication efficiency in federated learning through selective low-rank adaptation and differential privacy. Applied Sciences, 15(24), 13102. https://doi.org/10.3390/app152413102
16. Mosaiyebzadeh, F., Pouriyeh, S., Han, M., Liu, L., Xie, Y., Zhao, L., & Batista, D. M. (2025). Privacy-preserving federated learning-based intrusion detection system for IoHT devices. Electronics, 14(1), 67. https://doi.org/10.3390/electronics14010067
17. Rampone, G., Ivaniv, T., & Rampone, S. (2025). A hybrid federated learning framework for privacy-preserving near-real-time intrusion detection in IoT environments. Electronics, 14(7), 1430. https://doi.org/10.3390/electronics14071430
18. Salawu, G., & Glen, B. (2026). Integrating artificial intelligence into mechatronics: A comprehensive study of its influence on system performance, autonomy, and manufacturing efficiency. Technologies, 14(3), 143. https://doi.org/10.3390/technologies14030143
19. Song, S., Li, Y., Wan, J., Fu, X., & Jiang, J. (2024). Data quality-aware client selection in heterogeneous federated learning. Mathematics, 12(20), 3229. https://doi.org/10.3390/math12203229
20. Zhao, J., Yang, M., Zhang, R., Song, W., Zheng, J., Feng, J., & Matwin, S. (2022). Privacy-enhanced federated learning: A restrictively self-sampled and data-perturbed local differential privacy method. Electronics, 11(23), 4007. https://doi.org/10.3390/electronics11234007