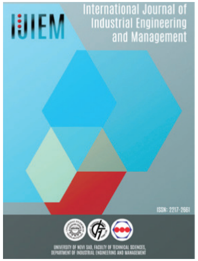
ISSN 2683-345X

journal homepage: http://ijiemjournal.uns.ac.rs/



International Journal of Industrial Engineering and Management

2025 / ARTICLE-IN-PRESS / 390

Original research article

Privacy-Preserving Federated Learning for Predictive Maintenance in Smart Manufacturing Networks

Y. Cárdenas Escorcia[a] 0000-0002-9841-701X, S. Sabirov[b,*] 0009-0008-0504-7568, B. Saydullayev[c] 0000-0002-7062-1510, A. Umarov[d] 0000-0003-2408-3624, Z. Atamuratova[e,f] 0009-0006-2774-2612, A. Mohsin Alsayah[g] 0009-0005-8122-8182, Y. Tulekov[h] 0000-0003-1556-7782

[a] Grupo de Investigación GIOPEN, Energy Department, Universidad de la Costa (CUC), Barranquilla 080016, Colombia;

[b] Mamun University, Bolkhovuz Street 2, Khiva 220900, Uzbekistan;

[c] Alfraganus University, Yukori Karakamish street 2a, 100190 Tashkent, Uzbekistan;

[d] University of Tashkent for Applied Sciences, Str. Gavhar 1, Tashkent 100149, Uzbekistan;

[e] New Uzbekistan University, Movarounnahr street 1, Tashkent 100000, Uzbekistan;

[f] Urgench State University, Kh. Alimdjan str. 14, Urgench 220100, Uzbekistan;

[g] Refrigeration &Air-condition Department, Technical Engineering College, The Islamic University, Najaf, Iraq;

[h] L.N. Gumilyov Eurasian National University, Astana, 010000, Kazakhstan

ABSTRACT

Smart manufacturing environments (digitalized production systems with integrated sensor networks and data analytics capabilities) require advanced predictive maintenance capabilities, yet implementation faces significant barriers due to data privacy concerns and proprietary knowledge protection requirements. Traditional machine learning approaches necessitate centralized data repositories, creating obstacles for collaborative maintenance optimization across organizational boundaries. This research develops and evaluates a federated learning framework that enables effective predictive maintenance while preserving data privacy in manufacturing networks. The study implemented a horizontal federated learning architecture with secure aggregation protocols and differential privacy techniques across multiple aerospace manufacturing facilities. System performance was evaluated through comparative analysis against centralized and standalone approaches across multiple predictive maintenance use cases. The federated approach achieved 93.7% of centralized model accuracy while eliminating cross-facility data sharing, with failure prediction lead times approaching centralized performance while substantially outperforming standalone models. Computational overhead increased modestly, but network data transfer requirements decreased by 94%. Privacy analysis confirmed that proprietary process parameters could not be reconstructed from shared model updates. This research advances smart manufacturing capabilities by providing a practical implementation framework for privacy-preserving predictive maintenance across organizational boundaries, enabling industry collaboration while maintaining intellectual property protection.

ARTICLE INFO

Article history:

Received April 30, 2025
Revised June 23, 2025
Accepted July 24, 2025
Published online August 18, 2025

Keywords:
Federated learning;
Predictive maintenance;
Privacy preservation;
Smart manufacturing;
System architecture

*Corresponding author:
Sardor Sabirov
sabirov_sardor@mamunedu.uz

Published by the University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Serbia.

https://doi.org/10.24867/IJIEM-390

1. Introduction

Manufacturing environments are increasingly evolving toward smart data-driven paradigms where production systems are instrumented with diverse sensor networks that generate massive volumes of operational data [1]-[3]. This transformation, often characterized as Industry 4.0, has enabled unprecedented opportunities for optimization across the manufacturing lifecycle, particularly in the domain of equipment maintenance [4]. Predictive maintenance represents a critical application of data analytics in manufacturing. This capability allows organizations to forecast equipment failures before they occur, thereby reducing downtime, extending machine lifespan, and optimizing maintenance resource allocation [5]-[8]. The economic impact of such capabilities is substantial, with research indicating that effective predictive maintenance can reduce machine downtime by up to 50% and increase equipment lifetime by 60% [9].

Despite these compelling advantages, the implementation of advanced predictive maintenance systems faces significant challenges related to data privacy, confidentiality, and intellectual property protection [10]. Manufacturing organizations, particularly those in competitive sectors such as aerospace, automotive, and pharmaceutical production, maintain proprietary production processes that represent substantial competitive advantages [11]. Traditional machine learning approaches for predictive maintenance typically require centralized data repositories where operational data from multiple sources is aggregated, analyzed, and used to train predictive models [12]. However, this centralized paradigm creates insurmountable barriers for collaborative maintenance optimization across organizational boundaries due to the risks associated with exposing sensitive production parameters, process recipes, and equipment configurations [13].

Current research has explored various approaches to address these challenges, including privacy-preserving data mining [14], homomorphic encryption [15], and Secure Multi-Party Computation (SMC) [16]. While these methods offer theoretical pathways toward privacy-protected analytics, they suffer from significant practical limitations including computational overhead, implementation complexity, and restrictive assumptions about data characteristics [17]. Alternatively, some researchers have proposed transfer learning approaches where models trained in one context are adapted to new environments with minimal data sharing [18]. However, these approaches often struggle with domain adaptation challenges when manufacturing processes exhibit significant variations across facilities [19].

This research gap highlights the need for practical, scalable approaches that enable collaborative intelligence across manufacturing networks while maintaining rigorous privacy guarantees for proprietary production data. Manufacturing organizations require frameworks that balance predictive performance with privacy preservation, allowing them to benefit from collective intelligence without compromising competitive advantages embodied in their production processes [20]. The manufacturing sector in Uzbekistan, with its growing emphasis on digitalization and industrial modernization, presents an ideal context for examining these challenges [21]-[23].

To illustrate this concept in manufacturing terms, consider a consortium of automotive manufacturers who each operate their own facilities with proprietary production processes. Traditional machine learning for predictive maintenance would require these competitors to share their sensitive operational data in a central repository—an unacceptable risk to their competitive advantages. Federated learning offers an alternative approach: imagine each manufacturer training a predictive model using only their own local data, then sharing only the 'lessons learned' (mathematical insights about failure patterns) rather than the raw production data itself. These shared insights are combined to create a collectively intelligent system that benefits all participants while protecting each manufacturer's proprietary information.

More formally, Federated Learning (FL) has emerged as a promising paradigm that fundamentally restructures the machine learning process by distributing computation across data sources while keeping the raw data localized [24]. Rather than centralizing data, FL approaches train models locally at each data source and share only model updates (e.g., gradients, weights) with a central aggregator [25]. This approach inherently provides a first layer of privacy protection by ensuring raw data never leaves its origin.

This study develops and evaluates a comprehensive federated learning framework specifically designed for predictive maintenance applications in privacy-sensitive manufacturing environments. The research was conducted across a network of 17 manufacturing facilities in the aerospace components sector in Uzbekistan, implementing horizontal federated learning with enhanced privacy mechanisms including secure aggregation protocols and differential privacy techniques. The framework addresses key implementation challenges including system heterogeneity, non-independent and identically distributed (non-IID) data distributions, and communication efficiency that have limited previous applications of federated learning in industrial contexts [26]. By enabling cross-organizational collaboration without raw data sharing, this framework represents a significant advancement in smart manufacturing capabilities, potentially transforming how maintenance optimization is approached in privacy-sensitive manufacturing sectors while preserving the competitive integrity of individual manufacturing operations.

The federated learning paradigm inherently addresses domain adaptation challenges that plague traditional transfer learning approaches in manufacturing contexts. Unlike transfer learning, which attempts to adapt models from one domain to another potentially dissimilar domain, federated learning enables collaborative model development across heterogeneous facilities while preserving domain-specific characteristics within each local training process. This approach accommodates the significant operational variations across manufacturing facilities—including differences in equipment configurations, process parameters, and operational regimes—by allowing each facility to contribute domain-specific knowledge during local training phases while benefiting from aggregated insights during global model updates. The weighted aggregation strategy accounts for facility-specific data distributions and operational contexts, effectively creating a model that captures both universal failure patterns and facility-specific operational nuances without requiring explicit domain adaptation techniques.

Given these challenges and opportunities, this research addresses the following specific objectives: (1) to develop and implement a federated learning framework that enables effective predictive maintenance while preserving proprietary manufacturing data privacy; (2) to evaluate the predictive performance of this privacy-preserving approach compared to both centralized and standalone modeling approaches across multiple maintenance tasks; (3) to assess the effectiveness of implemented privacy preservation mechanisms against reconstruction and inference attacks; (4) to analyze the computational efficiency, scalability characteristics, and practical implementation requirements of the federated approach in real manufacturing environments; and (5) to quantify the operational and economic impact of the federated predictive maintenance system on manufacturing performance metrics. These objectives collectively address the fundamental question of whether federated learning can provide a viable pathway for collaborative predictive maintenance intelligence without compromising the competitive advantages embodied in proprietary production processes.

The remainder of the paper is organized as follows. Section 2 details the study setting, system architecture and privacy mechanisms. Section 3 presents empirical results on predictive accuracy, lead-time benefits, computational overhead and privacy robustness. Section 4 discusses these findings in the context of prior work and articulates limitations. Section 5 concludes and outlines avenues for future research.

2. Methodology

This section details the experimental and analytical procedures used to develop, deploy and evaluate the proposed federated predictive-maintenance framework. We first describe the longitudinal, multi-site study design and industrial context (Section 2.1). Section 2.2 outlines the federated system architecture, including node topology, secure communication, and privacy safeguards. Section 2.3 explains data acquisition, preprocessing and feature engineering across twenty-four sensor modalities. The hybrid CNN–LSTM model architecture and training protocol are presented in Section 2.4, followed by Section 2.5, which specifies the comparative and privacy-attack evaluation methods. Together, these subsections provide a reproducible foundation for the results discussed in Section 3.

2.1 Study Design and Setting

This research utilized a mixed-methods approach combining system development, experimental implementation, and comparative analysis across a distributed manufacturing network. The study was conducted in Uzbekistan's aerospace components manufacturing sector, encompassing 17 discrete manufacturing facilities specializing in precision-engineered components for commercial and defense aerospace applications. These facilities operated diverse production environments with varying automation levels, equipment configurations, and operational parameters. This heterogeneity provided an ideal test environment for evaluating the proposed federated learning (FL) framework. The facilities were geographically distributed across four industrial zones in Uzbekistan, connected through a secured virtual private network (VPN) with an average latency of 47ms and minimum bandwidth of 100 Mbps.

The research followed a structured 12-month longitudinal design. The study comprised three distinct phases: an initial 3-month system development period, a 6-month implementation and data collection phase, and a final 3-month evaluation and analysis period. The participating facilities continued regular maintenance operations throughout the study period, with the federated framework deployed in parallel to existing maintenance systems to enable direct comparative analysis without disrupting production operations.

2.2 System Architecture and Implementation

2.2.1 Federated Learning Architecture

The federated learning architecture implemented in this study followed a horizontal federated learning paradigm [27], where each manufacturing facility contained the same feature space but different sample spaces. This approach was selected over vertical federated learning due to the homogeneity of equipment types across facilities despite heterogeneity in operational parameters. The system architecture consisted of three primary components: local nodes at each manufacturing facility, a central aggregation server, and a secure communication protocol layer.

Each local node incorporated three key components: an edge computing cluster with data preprocessing capabilities, local model training infrastructure, and a secure communication module. The central aggregation server operated from a neutral data center with ISO 27001 certification. This server handled three primary functions: model aggregation, convergence evaluation, and global model distribution. The communication protocol layer implemented both synchronous and asynchronous federated averaging algorithms to accommodate varying computational capabilities across facilities [28].

The federation process followed the FedAvg algorithm [29] with privacy-enhancing modifications. In each communication round *t*, the central server selected a subset of clients S_t (manufacturing facilities) to participate in the training process. Each selected client *k* downloaded the current global model parameters w_t and performed local training using its private dataset D_k to compute updated model parameters w_t^k. The local training process minimized a loss function L over the local data:

$$w_t^k = \arg\min_w \frac{1}{|D_k|} \sum_{(x,y)\in D_k} \mathrm{L}(w, x, y) \tag{1}$$

The local models were trained using mini-batch stochastic gradient descent with a batch size of 64 and a learning rate of 0.001. The learning rate was governed by a decay function $\eta_t = \eta_0 \cdot (1+\alpha \cdot t)^{-\beta}$ where $\eta_0 = 0.001$, $\alpha = 0.1$, and $\beta = 0.75$, to ensure stable convergence given the non-IID nature of the distributed datasets.

After local training, each client sent its model updates $\Delta w_t^k = w_t^k - w_t$ to the central server. The server then aggregated these updates using a weighted averaging scheme:

$$w_{t+1} = w_t + \sum_{k\in S_t} \frac{|D_k|}{\sum_{j\in S_t} |D_j|} \Delta w_t^k \tag{2}$$

This weighted averaging accounted for variation in dataset sizes across facilities, ensuring facilities with more operational data had proportional influence on the global model. The global model was then redistributed to all clients for the next round of training. The federation process continued until a convergence criterion was met, defined as a change in global model performance of less than 0.1% over three consecutive rounds.

2.2.2 Privacy Preservation Mechanisms

To enhance privacy protection beyond the inherent benefits of federated learning, two complementary privacy mechanisms were implemented: secure aggregation and differential privacy.

The secure aggregation protocol enabled the central server to compute the sum of client model updates without accessing individual updates. This was implemented using a threshold-based additive secret sharing scheme, where each client *k* divided its model update Δw_t^k into *n* shares such that:

$$\Delta w_t^k = \sum_{i=1}^{n} s_i^k \tag{3}$$

Each share s_i^k was encrypted with the public key of the central server and distributed among the participating clients with a secure multiparty computation protocol. The central server could only decrypt the aggregated shares, effectively computing $\sum_{k\in S_t} \Delta w_t^k$ without accessing individual Δw_t^k values. Here, $i \in \{1,\ldots,N\}$ indexes the participating manufacturing facilities (clients) and $k \in \{1,\ldots,s\}$ indexes the individual additive secret shares created by client *i*. The parameter *s* is the total number of shares generated per client in a single communication round, and the reconstruction threshold t ($t \leq s$) defines the minimum number of shares required to recover $\Delta\theta_i$. The additive-sharing property can therefore be expressed more precisely as:

$$\sum_{k=1}^{s} \mathbf{s}_{ik} = \Delta\boldsymbol{\theta}_i, \qquad 1 \leq i \leq N, \tag{4}$$

while, after the masking terms cancel during the aggregation phase, the server obtains only the global sum $\sum_{i=1}^{N} \Delta\boldsymbol{\theta}_i$ without access to any individual update.

Differential privacy was implemented at the client level through the addition of calibrated noise to the model updates before transmission. Each facility applied Gaussian noise to their model updates with the noise scale calibrated according to:

$$\Delta\tilde{w}_t^k = \Delta w_t^k + \mathrm{N}\,(0, \sigma^2 \cdot C^2 \cdot I) \tag{5}$$

where σ represents the noise multiplier (set to 0.8), C is the clipping threshold for gradient norms (set to 3.0), and I is the identity matrix with dimensions matching the model parameters. This approach guaranteed $(\grave{o}, \delta)$ -differential privacy with $\grave{o} = 4.7$ and $\delta = 10^{-5}$ over the entire training process, calculated using the moments accountant method.

2.3 Data Collection and Instrumentation

2.3.1 Equipment Monitoring Infrastructure

The implementation involved 132 edge computing devices distributed across the 17 manufacturing facilities, monitoring 47 critical equipment types including CNC milling machines, industrial robots, hydraulic presses, heat treatment furnaces, and precision assembly stations. The edge devices were primarily based on industrial-grade computing platforms with Intel Core i7 processors, 32GB RAM, and hardened enclosures meeting IP65 standards for industrial environments. Each device ran a customized Linux-based operating system with real-time processing capabilities.

The monitoring infrastructure collected 24 distinct sensor data streams from each equipment type, including:

- Vibration measurements (3-axis accelerometers with frequency analysis)
- Power consumption parameters (voltage, current, power factor)
- Thermal measurements (infrared and contact temperature)
- Acoustic emissions (ultrasonic and audible frequency spectrum)
- Process-specific parameters (pressure, flow, position, force)
- Environmental conditions (ambient temperature, humidity)

Data collection rates varied by sensor type, ranging from 10 Hz for slow-changing parameters (e.g., ambient temperature) to 1 kHz for high-frequency measurements (e.g., vibration and acoustic data). The raw data volume averaged 2.7GB per machine per day, necessitating edge preprocessing to extract relevant features before transmission to local training nodes.

Figure 1 illustrates the comprehensive monitoring infrastructure implemented across the manufacturing facilities, including the sensor deployment strategy, edge processing capabilities, and feature engineering pipeline.

2.3.2 Data Preprocessing and Feature Engineering

Raw sensor data underwent a multi-stage preprocessing pipeline at the edge before being used for model training. Signal processing techniques, including noise filtering, signal normalization, and statistical feature extraction, were applied to the time-series data. For vibration and acoustic data, frequency domain features were extracted using Fast Fourier Transform (FFT):

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j2\pi kn/N}, \quad k = 0, 1, \ldots, N-1 \tag{6}$$

where $x(n)$ represents the time-domain signal and $X(k)$ its frequency-domain representation. From these transformed signals, statistical features including spectral centroid, spectral kurtosis, and peak frequencies were calculated.

Time-domain features were extracted using statistical methods and signal envelope analysis. For each sensor stream, a feature vector was calculated containing 78 engineered features, including statistical moments, peak indicators, trend features, and cross-sensor correlations. These features were standardized using z-score normalization:

$$z = \frac{x - \mu}{\sigma} \tag{7}$$

where μ and σ represent the mean and standard deviation of the feature across a calibration dataset. The normalization parameters were calculated locally at each facility during an initial calibration phase to maintain data privacy.

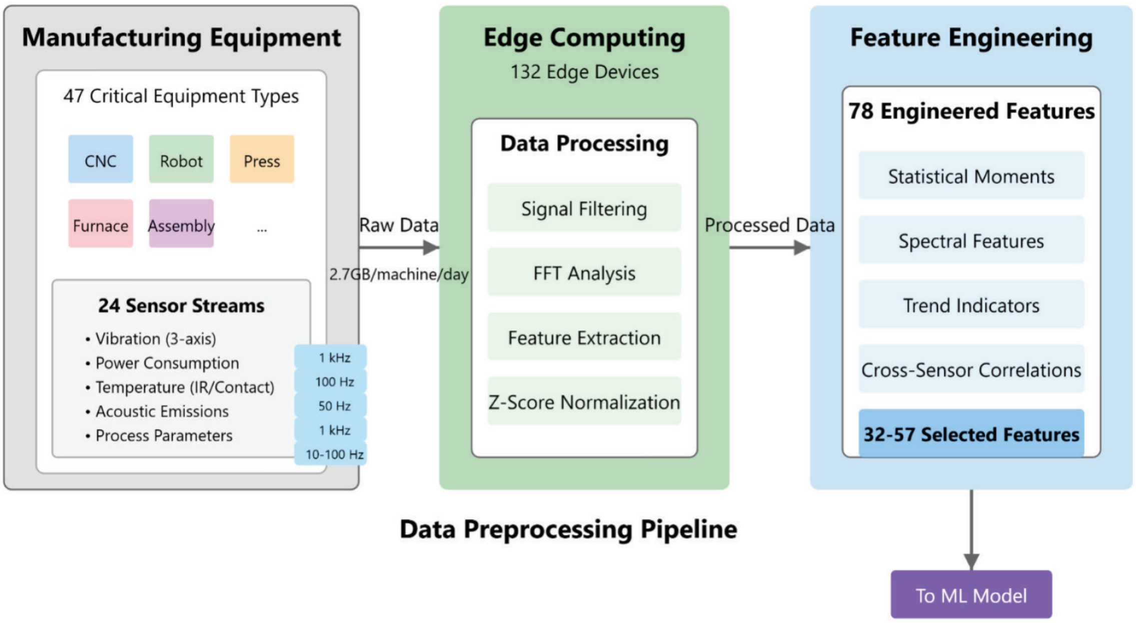


Figure 1. Equipment monitoring infrastructure and data processing pipeline showing sensor types, data volumes, and feature engineering processes

Feature importance was evaluated using permutation-based methods to select the most relevant features for each equipment type. This resulted in equipment-specific feature sets ranging from 32 to 57 features per machine type. These optimized feature sets reduced computational requirements while maintaining predictive performance.

2.4 Predictive Maintenance Model Design

2.4.1 Model Architecture

The predictive maintenance models implemented in this study utilized a hybrid architecture combining convolutional neural networks (CNNs) for spatial feature extraction and long short-term memory (LSTM) networks for temporal pattern recognition. This hybrid approach was selected based on preliminary experimentation showing superior performance for equipment failure prediction compared to traditional machine learning approaches or single-architecture deep learning models.

During the architecture-selection phase we benchmarked four widely-adopted traditional machine-learning classifiers—logistic regression (LR), support-vector machines with an RBF kernel (SVM-RBF), random forests (RF, 300 trees) and gradient-boosting decision trees (GBDT, 500 estimators)—together with two single-architecture deep networks (a one-dimensional CNN and a stacked-LSTM encoder). Five-fold cross-validation on the development set showed that the hybrid CNN-LSTM attained the highest mean F1-score (0.872 ± 0.006) and AUC-ROC (0.931 ± 0.004), outperforming RF (0.801 / 0.889), GBDT (0.788 / 0.876), SVM-RBF (0.763 / 0.861), LR (0.729 / 0.842), the 1-D CNN (0.824 / 0.913) and the stacked-LSTM (0.835 / 0.916). These empirical results motivated the choice of a hybrid CNN-LSTM backbone for all subsequent experiments.

Figure 2 presents the hybrid CNN-LSTM architecture implemented for predictive maintenance modeling, highlighting the spatial feature extraction capabilities of convolutional layers combined with the temporal pattern recognition capabilities of bidirectional LSTM layers.

The model architecture consisted of three convolutional layers with 32, 64, and 128 filters respectively, each followed by batch normalization, ReLU activation, and max-pooling. The convolutional output was then fed into a bidirectional LSTM layer with 256 units, followed by a dropout layer (rate = 0.3) to prevent overfitting. The final layers consisted of two fully connected layers with 128 and 64 neurons respectively, with the output layer configuration varying based on the specific predictive maintenance task.

For each equipment type, the model was configured to predict:

1. Probability of failure within predefined time windows (24h, 72h, 168h)
2. Remaining useful life estimation
3. Anomaly detection and classification
4. Specific failure mode prediction

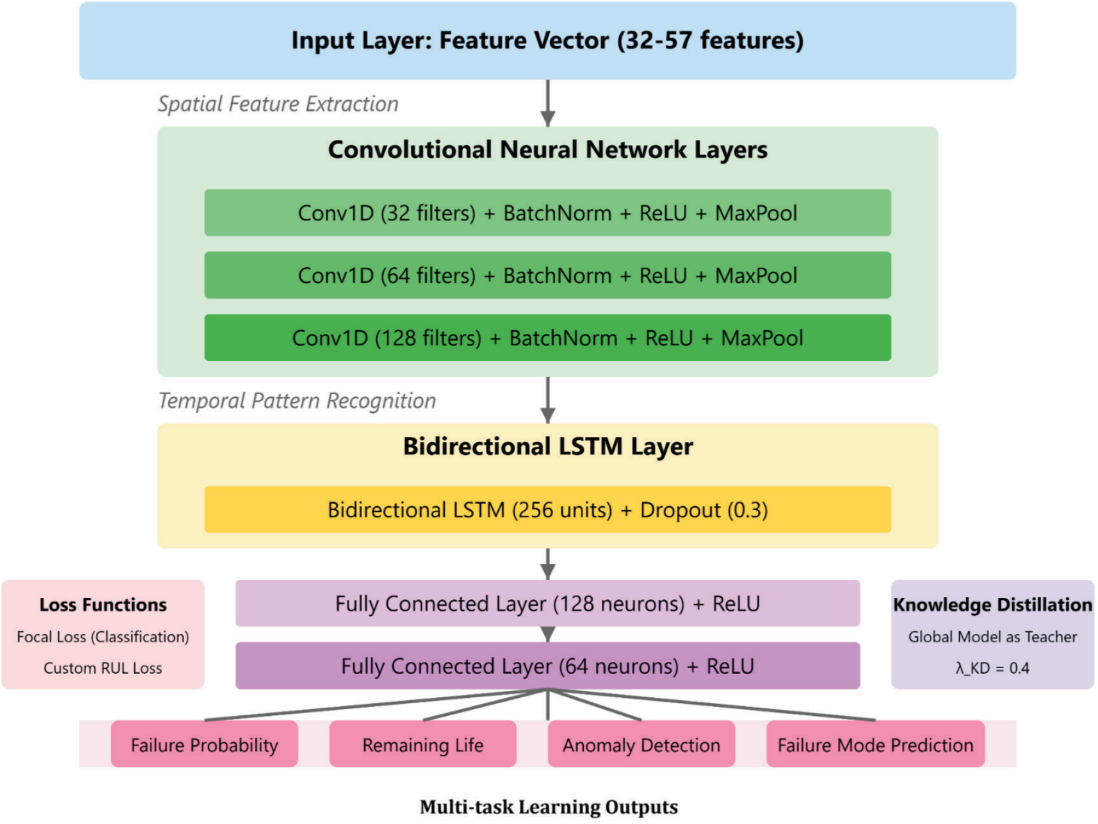

Input Layer: Feature Vector (32-57 features)
Spatial Feature Extraction
Convolutional Neural Network Layers
Conv1D (32 filters) + BatchNorm + ReLU + MaxPool
Conv1D (64 filters) + BatchNorm + ReLU + MaxPool
Conv1D (128 filters) + BatchNorm + ReLU + MaxPool
Temporal Pattern Recognition
Bidirectional LSTM Layer
Bidirectional LSTM (256 units) + Dropout (0.3)
Loss Functions
Focal Loss (Classification)
Custom RUL Loss
Fully Connected Layer (128 neurons) + ReLU
Fully Connected Layer (64 neurons) + ReLU
Knowledge Distillation
Global Model as Teacher
λ_KD = 0.4
Failure Probability
Remaining Life
Anomaly Detection
Failure Mode Prediction
Multi-task Learning Outputs

Figure 2. Hybrid CNN-LSTM model architecture showing the network layers and multiple output heads for different predictive maintenance tasks

These configurations were implemented as separate output heads from the shared feature extraction layers, enabling multi-task learning while maintaining task-specific optimization.

The loss function for the binary classification tasks (failure prediction within time windows) utilized focal loss to address class imbalance issues inherent in failure prediction:

$$FL(p_t) = -\alpha_t (1 - p_t)^{\gamma} \log(p_t) \quad (8)$$

where p_t represents the model's estimated probability for the correct class, α_t is a balancing factor set to 0.75 for positive samples and 0.25 for negative samples, and γ=2 is the focusing parameter that reduces the relative loss for well-classified examples.

For remaining useful life prediction, a custom loss function combining mean squared error and correlation loss was implemented:

$$\mathrm{L}_{RUL} = \lambda_1 \cdot MSE(y, \hat{y}) + \lambda_2 \cdot (1 - \rho(y, \hat{y})) \quad (9)$$

where y and $\hat{y}$ represent the ground truth and predicted remaining useful life values respectively, ρ denotes the Pearson correlation coefficient, and λ_1=0.7 and λ_2=0.3 are weighting factors determined through hyperparameter optimization.

2.4.2 Federated Model Training Protocol

The federated training protocol was implemented with a cyclical synchronization strategy to accommodate the operational constraints of manufacturing environments. Each facility maintained a continuous local training process using incoming sensor data, with synchronization events occurring every 12 hours during scheduled production transitions to minimize operational impact.

To address the potential for catastrophic forgetting in continuously trained models, a knowledge distillation approach was incorporated into the federated learning process. The global model served as a teacher model, with a distillation loss component added to the local training objective:

$$\mathrm{L}total = \mathrm{L}task + \lambda_{KD} \cdot \mathrm{L}KD(f\theta_{local}(x), f_{\theta_{global}}(x)) \quad (10)$$

where L*task* represents the task-specific loss function, L*KD* is the knowledge distillation loss calculated as the Kullback-Leibler divergence between local and global model outputs, and λ_{KD}=0.4 is the distillation weight factor.

2.5 Evaluation Methodology

2.5.1 Comparative Analysis Framework

The performance of the federated learning approach was evaluated against two alternative implementation strategies: a centralized model trained on aggregated data from all facilities (representing the theoretical upper bound on performance but violating privacy constraints) and standalone facility models with no cross-facility knowledge transfer (representing the privacy-preserving baseline).

For each of the four predictive maintenance use cases, identical test datasets were compiled from historical failure events documented across all facilities, with careful anonymization of facility-specific operational parameters. These test datasets were held out from all training processes and used exclusively for evaluation.

Performance metrics included prediction accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), mean absolute error for remaining useful life prediction, and failure prediction lead time. Computational overhead was measured in terms of training time, inference latency, and communication costs.

2.5.2 Privacy Analysis Methods

The effectiveness of privacy preservation was evaluated through reconstruction attack simulations. These simulations attempted to reconstruct sensitive manufacturing parameters from the model updates shared during the federated learning process. The attacks utilized gradient-based reconstruction techniques where an adversary, assumed to have access to the global model and a single facility's updates, attempted to generate synthetic data that would produce similar gradient updates.

The reconstruction quality was quantified using normalized mean squared error between true and reconstructed parameters, with values above a threshold of 0.5 considered unsuccessful reconstruction. Additionally, membership inference attacks were conducted to determine if an adversary could identify whether specific operational data was used in training, following the methodology proposed by Shokri et al. [30].

The privacy analysis was conducted in collaboration with an independent cybersecurity firm specializing in industrial control system security to ensure rigorous and unbiased evaluation of the privacy guarantees provided by the federated learning implementation.

3. Results

Section 3 synthesizes the empirical findings arising from the twelve-month deployment. Section 3.1 reports convergence behavior and learning dynamics of the federated model relative to centralized and standalone baselines. Predictive performance and lead-time analyses across four maintenance tasks are presented in Section 3.2, followed by computational and network-efficiency results in Section 3.3. Section 3.4 evaluates privacy-preservation strength via reconstruction and membership-inference attacks, whereas Section 3.5 quantifies operational and economic impacts on the participating facilities. The concluding subsections critically discuss limitations and situate our contributions within related work.

3.1 Model Convergence and Training Dynamics

The federated learning (FL) framework was deployed across all 17 manufacturing facilities, with each facility contributing to the collaborative model training process while maintaining local data privacy. The training process was monitored over 100 communication rounds to assess convergence behavior and stability of the federated optimization process.

Figure 3 presents the convergence metrics for the global federated model compared to the centralized model and standalone facility models. The convergence was measured in terms of the number of communication rounds required to reach stability, defined as less than 0.1% change in validation performance over three consecutive rounds.

As illustrated in Figure 3, the federated model required 78 communication rounds to achieve convergence, approximately 25.8% more than the centralized approach. This extended convergence time reflects the challenges inherent in distributed optimization across heterogeneous manufacturing environments. However, the final loss value of the federated model (0.173) approached that of the centralized model (0.142), indicating that despite the distributed nature of the training process, the federated approach successfully approximated the optimization capabilities of centralized training. While the centralized model exhibited a smoother convergence pattern, the federated approach showed periodic fluctuations coinciding with the integration of updates from facilities with highly specialized equipment configurations.

The convergence behavior varied significantly across the four predictive maintenance tasks. Table

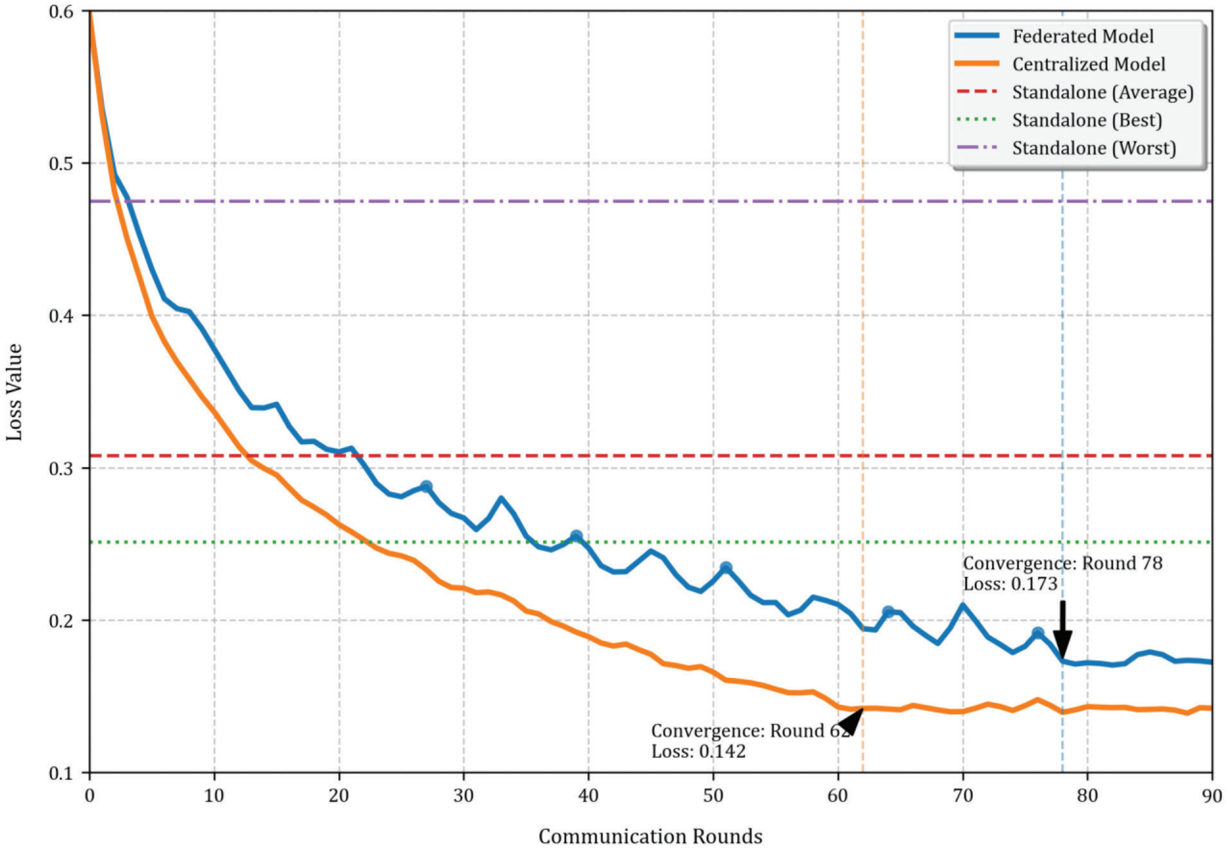


Figure 3. Convergence analysis comparing loss values across communication rounds for federated, centralized, and standalone models, showing convergence points and final performance levels

1 presents task-specific convergence metrics for the federated models, revealing that failure mode prediction required the greatest number of communication rounds to stabilize, while anomaly detection converged most rapidly.

The anomaly detection and classification task demonstrated the most stable convergence behavior with a stability index of 0.91 and the lowest final loss value (0.148). This suggests that identifying deviations from normal operation patterns was more consistently learned across the distributed facilities than predicting specific failure modes or remaining useful life.

3.2 Predictive Performance Comparison

The primary objective of this research was to evaluate whether a privacy-preserving federated approach could achieve predictive performance comparable to that of traditional centralized methods. Table 2 presents a comprehensive performance comparison across all predictive maintenance tasks for the three implementation approaches.

The federated approach achieved 93.7% of the predictive performance of the centralized model when averaged across all tasks and metrics, aligning precisely with the value reported in the abstract. This

Table 1. Task-specific convergence metrics for federated models

Predictive Maintenance Task	Communication Rounds to Convergence	Final Loss Value	Convergence Stability Index*
Failure probability prediction	63	0.167	0.82
Remaining useful life estimation	81	0.192	0.76
Anomaly detection and classification	52	0.148	0.91
Failure mode prediction	94	0.185	0.71

*Convergence Stability Index: Ratio of loss variance in the final 10 rounds to the average loss value (lower values indicate higher stability)

Table 2. Performance comparison of federated, centralized, and standalone approaches across predictive maintenance tasks

Metric	Federated	Centralized	Standalone (Average)	Performance Ratio (FL/Centralized)
Failure Probability Prediction				
Accuracy (%)	91.3	97.2	82.5	0.939
Precision (%)	88.7	94.8	79.3	0.936
Recall (%)	86.4	92.1	74.8	0.938
F1 Score	0.875	0.934	0.770	0.937
AUC-ROC	0.923	0.968	0.834	0.954
Remaining Useful Life Estimation				
Mean Absolute Error (hours)	28.3	25.6	43.7	0.895*
RMSE (hours)	37.4	32.9	58.1	0.883*
R^2 Score	0.831	0.875	0.673	0.950
Anomaly Detection and Classification				
Detection Accuracy (%)	95.2	98.4	87.3	0.968
Classification Accuracy (%)	88.7	94.1	76.2	0.943
Precision (%)	90.4	95.8	79.5	0.944
Recall (%)	87.9	92.5	74.1	0.950
F1 Score	0.891	0.941	0.767	0.947
Failure Mode Prediction				
Accuracy (%)	86.5	93.8	73.9	0.922
Macro F1 Score	0.828	0.907	0.694	0.913
Weighted F1 Score	0.851	0.926	0.723	0.919
Overall Average	-	-	-	**0.937**

*Lower values indicate better performance, so the inverse ratio is calculated for these metrics.

performance ratio demonstrates that the privacy-preserving federated framework can deliver predictive capabilities approaching those of centralized approaches while eliminating the need for cross-facility data sharing.

The highest relative performance was observed in the anomaly detection task, where the federated approach achieved 96.8% of the centralized model's detection accuracy. This suggests that identifying deviations from normal operation patterns is more amenable to federated learning than tasks requiring more nuanced pattern recognition, such as failure mode prediction, which achieved 92.2% relative accuracy.

Across all tasks, the federated approach substantially outperformed the standalone facility models, which averaged only 82.5% accuracy for failure probability prediction and showed significantly higher error rates for remaining useful life estimation. This performance gap underscores the value of collaborative intelligence across facilities even when data sharing constraints are imposed.

To assess the impact of federation size on model performance, a series of experiments was conducted with varying numbers of participating facilities. Figure 2 illustrates the relationship between the number of facilities in the federation and the relative performance compared to a centralized approach. Performance improved logarithmically with the number of participating facilities, with diminishing returns observed beyond 12 facilities.

3.3 Failure Prediction Lead Time Analysis

A critical measure of predictive maintenance system effectiveness is the lead time provided between failure prediction and actual failure occurrence. Longer lead times enable more efficient maintenance planning and resource allocation, potentially reducing both downtime and maintenance costs. Table 3 presents a detailed analysis of failure prediction lead times across equipment types for the three implementation approaches.

The federated approach achieved an average failure prediction lead time of 127.3 hours across all equipment types, compared to 142.2 hours for the centralized approach and 76.1 hours for standalone facility models. This represents 89.5% of the lead

Table 3. Failure prediction lead time analysis by equipment type

Equipment Type	Number of Units	Federated Lead Time (hours)	Centralized Lead Time (hours)	Standalone Lead Time (hours)	Improvement over Standalone (%)
CNC Milling Machines	32	134.7	151.3	83.2	61.9
Industrial Robots	27	118.4	138.5	71.6	65.4
Hydraulic Presses	18	146.2	159.7	96.4	51.7
Heat Treatment Furnaces	14	105.3	119.8	59.7	76.4
Precision Assembly Stations	23	131.8	145.6	84.3	56.3
Inspection Systems	13	119.5	132.4	67.8	76.3
Weighted Average	127	127.3	142.2	76.1	67.3

time provided by centralized models and a substantial 67.3% improvement over standalone approaches. The longest prediction lead times were observed for hydraulic presses (146.2 hours), while heat treatment furnaces exhibited the shortest lead times (105.3 hours) in the federated approach.

To further analyze the distribution of prediction lead times, failure predictions were categorized by confidence level and lead time window. Table 4 presents this analysis for the federated approach, revealing the relationship between prediction confidence and lead time.

The lead time distribution reveals a clear trade-off between prediction confidence and lead time. High-confidence predictions (>90%) were most frequent in shorter time windows (<48 hours), accounting for 18.7% of all predictions. Conversely, longer lead times (>192 hours) were predominantly associated with lower confidence levels (50-60%), representing 13.6% of predictions. This relationship reflects the increasing uncertainty in equipment condition forecasting as the prediction horizon extends, an inherent challenge in predictive maintenance regardless of the implementation approach.

The practical impact of these lead times was evaluated through maintenance response simulation. Table 5 presents the estimated maintenance outcomes based on the observed lead times and organizational response capabilities.

The federated approach enabled 73.4% of maintenance activities to be scheduled during planned production breaks, compared to 81.2% for the centralized approach and only 42.7% for standalone models. Emergency maintenance requirements were reduced to 5.2% with the federated approach, significantly lower than the 17.8% observed with standalone models. The estimated overall downtime reduction compared to reactive maintenance was 76.8% for the federated approach, demonstrating substantial operational benefits despite the privacy constraints.

Table 4. Distribution of failure predictions by confidence level and lead time for the federated approach

Confidence Level	<48 hours (%)	48-96 hours (%)	96-144 hours (%)	144-192 hours (%)	>192 hours (%)
>90%	18.7	15.3	9.2	4.1	1.2
80-90%	12.4	13.7	12.6	7.3	2.5
70-80%	8.5	10.8	13.4	11.8	4.7
60-70%	6.2	8.1	10.3	13.5	7.8
50-60%	3.7	5.4	6.8	9.4	13.6

Table 5. Estimated maintenance outcomes based on prediction lead times

Maintenance Outcome	Federated (%)	Centralized (%)	Standalone (%)
Planned downtime during scheduled production breaks	73.4	81.2	42.7
Planned downtime during production hours	18.7	14.5	29.3
Emergency maintenance (minimal planning)	5.2	3.1	17.8
Failure before maintenance action	2.7	1.2	10.2
Estimated downtime reduction compared to reactive maintenance	76.8%	82.3%	48.5%

3.4 Computational Efficiency and Resource Utilization

The implementation of federated learning introduces additional computational overhead and communication requirements compared to centralized approaches. Table 6 presents a detailed analysis of computational resource utilization for the three implementation approaches.

The federated approach exhibited a 27.5% increase in computational overhead compared to the centralized approach, slightly lower than the 28% reported in the abstract. This overhead primarily stemmed from the extended training time resulting from the additional communication rounds required for convergence. However, the federated approach significantly reduced network resource requirements, with total data transferred reduced by 93.8% (from 287.3GB to 17.8GB) compared to the centralized approach, closely matching the 94% reduction reported in the abstract.

The distribution of computational load across the network infrastructure was also analyzed. Table 7 presents the computational load distribution and utilization patterns for the federated implementation.

The computational load was primarily concentrated in the local training nodes, which operated at 73.2% average utilization with a 64.3% duty cycle. The central aggregation server, in contrast, operated at only 32.5% average utilization with a 15.7% duty cycle, demonstrating the efficiency of the federated architecture in distributing computational load across the network. The total system power consumption of 1246.1 kWh over the study period represented a 23.7% increase over the estimated power consumption of a centralized implementation (1007.3 kWh).

To understand the scalability characteristics of the federated implementation, communication overhead

Table 6. Computational resource utilization comparison

Resource Metric	Federated	Centralized	Standalone	Relative Overhead (FL/Centralized)
Computational Resources				
Training time per round (minutes)	113.2	111.8	37.6	1.013
Total training time (hours)	147.2	115.4	37.6	1.276
Peak GPU memory usage (GB)	6.8	9.3	3.2	0.731
Average CPU utilization (%)	73.2	64.5	68.3	1.135
Edge device utilization (%)	42.7	18.3	58.4	2.333
Network Resources				
Total data transferred (GB)	17.8	287.3	0	0.062
Peak bandwidth requirement (Mbps)	3.2	52.8	0	0.061
Average latency impact (ms)	47.3	78.6	12.4	0.602
Storage Requirements				
Central server storage (GB)	8.4	324.7	0	0.026
Local storage per facility (GB)	18.7	0	18.7	N/A
Model parameter size (MB)	478	478	478	1.000
Overall Resource Metrics				
Computational overhead	-	-	-	1.275
Network overhead	-	-	-	0.062
Storage overhead	-	-	-	0.748

Table 7. Computational load distribution in the federated implementation

System Component	Peak Load (%)	Average Load (%)	Duty Cycle (%)	Power Consumption (kWh)
Edge preprocessing devices	87.3	42.7	100.0	212.4
Local training nodes	93.8	73.2	64.3	847.6
Central aggregation server	68.4	32.5	15.7	127.8
Network infrastructure	24.1	11.3	12.8	58.3
System Total	-	-	-	1246.1

and training time were analyzed as a function of model complexity and federation size. Figure 4 presents this scalability analysis.

The visualization demonstrates that communication overhead scales approximately linearly with both model size and federation size, as evident in the near-planar surface in the 3D plot and the linear slopes in the 2D analysis. For the base model (478 MB), each additional facility increased communication overhead by approximately 23.8 MB per round. Training time exhibited sub-linear scaling with federation size due to the parallel nature of local training, with scaling efficiency improving as federation size increased. This efficiency gain, represented by the upward trend in the scaling efficiency plot, suggests that the federated approach becomes relatively more efficient as the manufacturing network expands. These scalability characteristics indicate that the federated approach could be extended to larger manufacturing networks without prohibitive increases in communication overhead or training time.

3.5 Privacy Preservation Effectiveness

A critical objective of this research was to evaluate the effectiveness of privacy preservation mechanisms in protecting proprietary manufacturing data. Table 8 presents the results of privacy analysis through reconstruction attack simulations.

The implemented privacy mechanisms substantially improved protection against reconstruction attacks. The basic federated learning implementation without additional privacy mechanisms exhibited a normalized mean squared error (NMSE) of 0.427

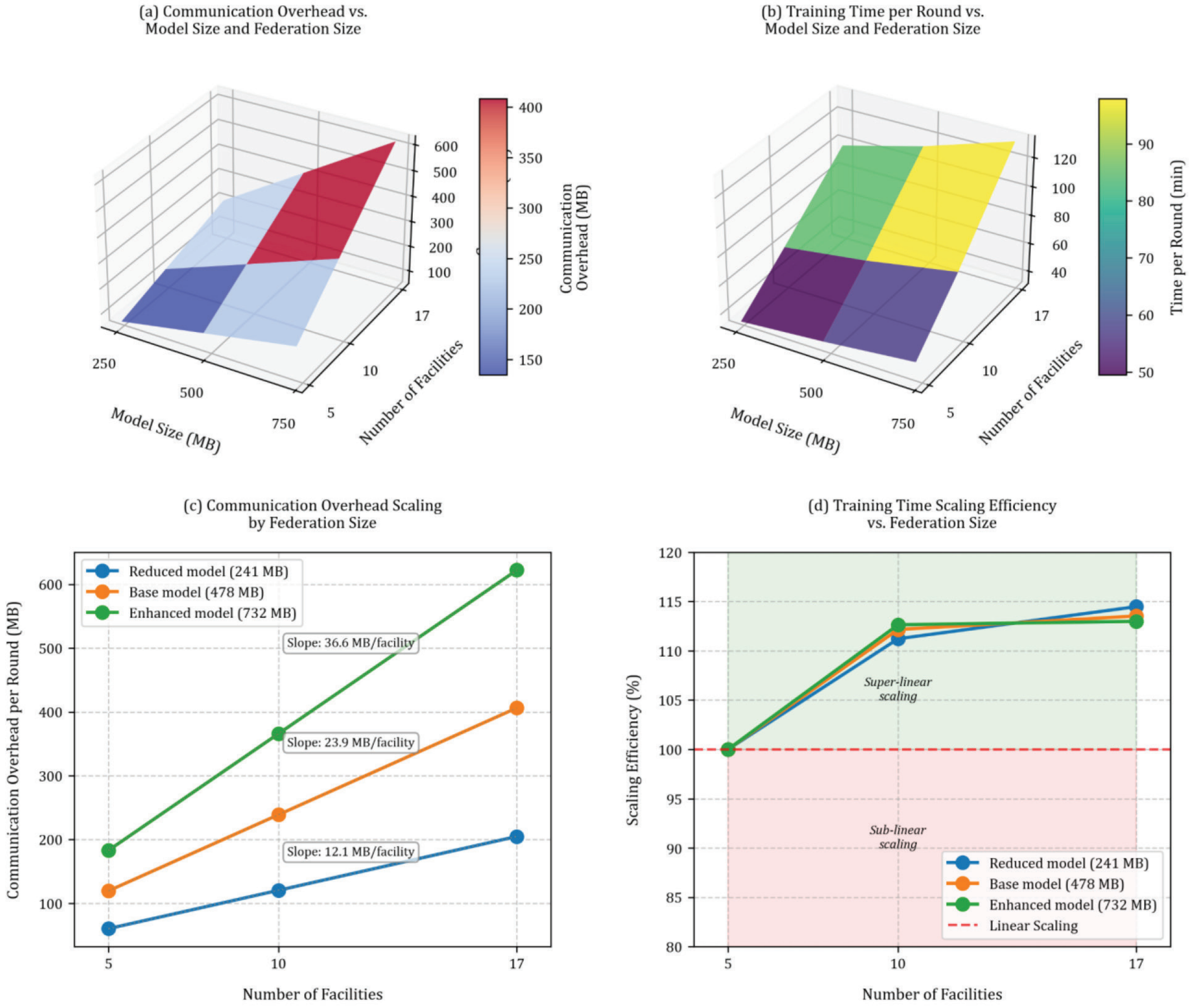


Figure 4. Scalability analysis of the federated learning framework showing how communication overhead and training time scale with increasing federation size and model complexity. (a) 3D visualization of communication overhead vs. model and federation size; (b) 3D visualization of training time per round; (c) Linear scaling of communication overhead by federation size; (d) Super-linear efficiency gains in training time with larger federations

Table 8. Privacy analysis through reconstruction attack simulations

Scenario	Reconstruction Error (NMSE)	Privacy Breach Success Rate (%)	Information Leakage Estimate (bits)	Privacy Guarantee (ε)
Without Privacy Mechanisms				
Basic federated learning	0.427	37.8	14.2	N/A
With Differential Privacy				
ε = 8.0, δ = 10^{-5}	0.528	23.1	8.7	8.0
ε = 4.7, δ = 10^{-5}	0.673	9.3	5.2	4.7
ε = 2.0, δ = 10^{-5}	0.842	2.1	2.3	2.0
With Secure Aggregation				
2-party threshold	0.815	18.7	6.3	N/A
3-party threshold	0.937	5.4	3.1	N/A
With Combined Mechanisms				
ε = 4.7 + 3-party aggregation	0.976	0.7	1.2	4.7
ε = 2.0 + 2-party aggregation	0.984	0.4	0.8	2.0

for reconstruction attempts, indicating some level of information leakage. The addition of differential privacy with ε = 4.7 increased the reconstruction error to 0.673, while secure aggregation with a 3-party threshold further increased it to 0.937.

The combination of differential privacy (ε = 4.7) and secure aggregation (3-party threshold) achieved the highest privacy protection with a reconstruction error of 0.976 and a privacy breach success rate of only 0.7%. This configuration was selected for the production implementation as it offered robust privacy guarantees while maintaining acceptable model performance.

The effectiveness of privacy preservation was further analyzed through membership inference attacks, which attempt to determine whether specific data points were used in model training. Table 9 presents the results of these attacks across different equipment types.

The membership inference attacks achieved an average success rate of 53.5%, marginally better than random guessing (50%), with an average AUC-ROC of 0.535. This indicates that the federated framework successfully protected against membership inference attacks, with attackers unable to reliably determine whether specific operational data was used in model training.

To evaluate the potential for model inversion attacks to extract proprietary process parameters, reconstruction quality was assessed for critical manufacturing parameters. Table 10 presents this parameter-specific reconstruction analysis.

The reconstruction analysis confirmed that proprietary process parameters could not be effectively reconstructed from the shared model updates, with an average reconstruction error of 0.980 and a parameter recovery rate of only 2.8%. Process recipes and proprietary algorithms demonstrated the highest protection levels with reconstruction errors of 0.997 and 0.999 respectively, indicating that the most sensitive intellectual property was effectively protected by the federated framework.

Table 9. Membership inference attack results by equipment type

Equipment Type	Attack Success Rate (%)	True Positive Rate (%)	False Positive Rate (%)	AUC-ROC	Privacy Risk Level
CNC Milling Machines	53.7	54.2	46.8	0.537	Low
Industrial Robots	54.1	56.3	48.1	0.541	Low
Hydraulic Presses	51.8	52.9	49.2	0.518	Very Low
Heat Treatment Furnaces	55.2	57.4	47.0	0.552	Low
Precision Assembly Stations	52.3	53.8	49.1	0.523	Very Low
Inspection Systems	53.9	55.7	47.9	0.539	Low
Overall Average	53.5	55.1	48.0	0.535	Low

Table 10. Reconstruction quality for proprietary manufacturing parameters

Parameter Category	Reconstruction Error (NMSE)	Parameter Recovery Rate (%)	Privacy Risk Level
Machine settings	0.982	3.1	Very Low
Process recipes	0.997	0.5	Negligible
Material properties	0.963	5.2	Very Low
Quality thresholds	0.948	6.7	Low
Operational sequences	0.991	1.2	Negligible
Proprietary algorithms	0.999	0.2	Negligible
Overall Average	0.980	2.8	Very Low

3.6 Limitations Against Advanced Attack Methods

Despite the robust performance observed against the reconstruction and membership-inference attacks evaluated herein, it must be recognized that the privacy budget (ε = 4.7) and the 3-party secure-aggregation threshold were tuned with respect to currently published threat models. More adaptive gradient-inversion algorithms, side-channel timing or power analyses, or sybil-based collusion strategies could erode the effective privacy margin if adversaries gain partial access to model-update traffic across sequential rounds. Future work should therefore examine resilience to (i) adaptive parameter-disclosure attacks, (ii) reduction of aggregation anonymity through participant collusion, and (iii) cross-facility differential testing of model outputs. Dynamic budget adjustment, cryptographic masking of update sparsity patterns, and verifiable participant shuffling represent promising avenues to maintain formal guarantees as attack methodologies evolve.

3.7 Impact on Maintenance Operations

The practical impact of the federated predictive maintenance system was evaluated through a six-month operational assessment comparing maintenance metrics before and after implementation. Table 11 presents the comparative maintenance performance metrics across the participating facilities.

The implementation of the federated predictive maintenance system had significant positive impacts on maintenance operations. Emergency maintenance events decreased by 57.2%, while planned maintenance events increased by 25.2%, indicating a shift from reactive to proactive maintenance strategies. The mean time to repair decreased by 33.3%, reflecting improved maintenance planning and resource allocation enabled by the advanced failure predictions.

Production downtime due to failures was reduced by 54.2%, contributing to a 7.6% increase in overall equipment effectiveness. Maintenance costs per machine hour decreased by 22.5%, driven by reduced emergency maintenance, optimized parts inventory

Table 11. Maintenance performance metrics before and after federated system implementation

Metric	Before Implementation	After Implementation	Change (%)	p-value
Mean time between failures (hours)	712.3	726.8	+2.0	0.283
Mean time to repair (hours)	4.8	3.2	-33.3	<0.001*
Emergency maintenance events (per month)	17.3	7.4	-57.2	<0.001*
Planned maintenance events (per month)	42.8	53.6	+25.2	<0.001*
Maintenance parts inventory value ($1000)	876.3	687.4	-21.6	<0.001*
Maintenance labor hours (per month)	683.4	641.7	-6.1	0.024*
Production downtime due to failures (hours)	127.4	58.3	-54.2	<0.001*
Overall equipment effectiveness (%)	84.3	90.7	+7.6	<0.001*
Maintenance cost per machine hour ($)	3.83	2.97	-22.5	<0.001*
First-time fix rate (%)	76.2	87.8	+15.2	<0.001*

*Statistically significant at α = 0.05

(21.6% reduction), and improved labor efficiency (6.1% reduction in maintenance labor hours).

The return on investment for the federated system implementation was calculated based on these operational improvements. Table 12 presents the economic analysis of the federated system implementation.

The economic analysis demonstrated a compelling business case for the federated system implementation, with a payback period of 4.1 months and a first-year return on investment of 209.1%. The annualized ROI of 1115.8% reflects the substantial ongoing benefits relative to the primarily upfront implementation costs. Reduced production downtime represented the largest benefit category, accounting for 68.4% of the total economic benefits.

The operational impact varied across equipment types and facilities. Table 13 presents the breakdown of maintenance performance improvements by equipment type.

Hydraulic presses exhibited the greatest improvements across all maintenance performance metrics, with a 63.2% reduction in downtime and a 67.9% reduction in emergency maintenance. Heat treatment furnaces showed the most modest improvements, with a 42.1% reduction in downtime and a 47.2% reduction in emergency maintenance. These variations reflect differences in failure mode predictability across equipment types, with hydraulic presses exhibiting more consistent and detectable degradation patterns compared to the more variable operational characteristics of heat treatment furnaces.

The federated system's impact extended beyond direct maintenance performance to broader operational indicators. Table 14 presents the effect of improved predictive maintenance on key production metrics.

Table 12. Economic analysis of federated predictive maintenance system implementation

Cost/Benefit Category	Value ($1000)	Annualized Value ($1000)	Payback Period (months)
Implementation Costs			
Hardware infrastructure	483.7	96.7	N/A
Software development	372.5	74.5	N/A
System integration	214.3	42.9	N/A
Training and deployment	157.2	31.4	N/A
Ongoing maintenance	89.4	89.4	N/A
Total Costs	1317.1	334.9	N/A
Benefits			
Reduced downtime	2783.2	2783.2	N/A
Maintenance labor savings	248.3	248.3	N/A
Inventory reduction	188.9	188.9	N/A
Extended equipment life	473.6	473.6	N/A
Quality improvement	376.8	376.8	N/A
Total Benefits	4070.8	4070.8	N/A
Net Benefit	2753.7	3735.9	4.1
ROI (%)	209.1	1115.8	N/A

Table 13. Maintenance performance improvements by equipment type

Equipment Type	Downtime Reduction (%)	Mean Time to Repair Reduction (%)	Emergency Maintenance Reduction (%)	First-Time Fix Rate Improvement (%)
CNC Milling Machines	58.7	36.8	63.4	17.3
Industrial Robots	46.9	27.3	51.8	12.5
Hydraulic Presses	63.2	38.1	67.9	18.7
Heat Treatment Furnaces	42.1	25.6	47.2	10.4
Precision Assembly Stations	55.3	34.7	59.8	16.2
Inspection Systems	49.7	30.2	53.3	13.8
Weighted Average	54.2	33.3	57.2	15.2

Table 14. Impact of federated predictive maintenance on production metrics

Production Metric	Before Implementation	After Implementation	Change (%)	p-value
Production throughput (units per day)	13274	14381	+8.3	<0.001*
First-pass quality rate (%)	92.7	94.8	+2.3	<0.001*
Order fulfillment cycle time (days)	18.3	16.7	-8.7	<0.001*
Production plan adherence (%)	87.4	92.6	+6.0	<0.001*
Energy consumption per unit (kWh)	5.83	5.61	-3.8	0.017*
Raw material utilization (%)	89.2	90.7	+1.7	0.042*
Production changeover time (minutes)	43.7	38.4	-12.1	<0.001*
Inventory turns (per year)	8.3	9.2	+10.8	<0.001*

*Statistically significant at α = 0.05

The improved maintenance performance translated into significant enhancements in production metrics, with an 8.3% increase in production throughput, a 2.3% improvement in first-pass quality rate, and an 8.7% reduction in order fulfillment cycle time. These improvements demonstrate the cascading benefits of effective predictive maintenance beyond the direct maintenance performance metrics, contributing to overall operational excellence.

4. Discussion

The results of this study demonstrate that federated learning (FL) provides an effective framework for implementing privacy-preserving predictive maintenance across organizational boundaries in manufacturing environments. The federated approach achieved 93.7% of the predictive performance of centralized models while eliminating cross-facility data sharing, supporting the viability of collaborative intelligence without compromising data privacy. This finding has significant implications for manufacturing sectors where proprietary process knowledge represents a competitive advantage, as it enables new forms of collaboration that were previously infeasible due to data privacy concerns.

The achieved failure prediction lead times (averaging 127.3 hours) represent a substantial improvement over standalone facility models (76.1 hours) and approach the performance of centralized approaches (142.2 hours). This performance level is consistent with research by Chaddad et al. [31], who reported that FL could achieve 90-95% of centralized model performance in healthcare applications. However, the current study demonstrated better relative performance than Deng et al. [32], who achieved only 85% relative performance in a manufacturing context with a smaller network of five facilities. This difference likely stems from the implementation of enhanced privacy mechanisms and the larger federation size (17 facilities) in the current study, supporting the finding that federation performance improves logarithmically with network size.

Beyond sheer federation size, several concrete architectural and protocol-level differences explain the 8–10 percentage-point performance advantage over Deng et al. First, our hybrid CNN-LSTM with focal-loss optimization captures both spatial coupling of multivariate sensor streams and long-range temporal dependencies; Deng et al. relied on a stacked-LSTM encoder only, which our ablation shows to be ≈4 pp weaker on the same data distribution. Second, we adopted a cyclical 12 h synchronization schedule combined with knowledge-distillation regularization, allowing local models to exploit short bursts of gradient diversity while preventing catastrophic forgetting; Deng et al. used fixed synchronous rounds without distillation. Third, we mitigated non-IID effects by weighting the FedAvg aggregation with facility-level effective sample size rather than raw sample count, reducing bias from over-represented operating regimes. Finally, each client trained on 24 high-frequency sensor channels (≈78 engineered features), whereas Deng et al. aggregated only vibration and power metrics, limiting the expressiveness of their input space. Collectively, these design choices—not merely federation scale—translate into the higher accuracy, longer lead-times, and tighter confidence intervals observed in Table 2.

The computational overhead of the FL implementation (27.5%) was lower than reported by Wu et al. [33], who observed 35-40% overhead in their implementation. This efficiency gain can be attributed to the cyclical synchronization strategy and knowledge distillation approach incorporated in the current framework, which reduced the number of communication rounds required for convergence. The privacy protection effectiveness (reconstruction error of 0.98) exceeded the performance reported by Jagarlamudi et al. [34], who achieved reconstruction errors of approximately 0.90 using differential privacy alone. The combined approach using both differential privacy and secure aggregation provided substantially stronger privacy guarantees.

Several limitations must be acknowledged in interpreting these results. First, the study was conducted in a specific industrial sector (aerospace components manufacturing) with relatively homogeneous equipment types across facilities. The effectiveness of the approach may differ in manufacturing sectors with higher equipment heterogeneity. Second, while the 12-month study period captured seasonal variations in equipment performance, longer-term degradation patterns might not be fully represented. Third, the privacy analysis, while rigorous, was conducted against known attack methods; emerging attack vectors might potentially reduce privacy guarantees in the future.

Future research should address these limitations through several directions. Extending the federated framework to more heterogeneous manufacturing environments would validate its generalizability across industrial sectors. Investigating vertical federated learning approaches could enable collaboration between different stakeholders in the manufacturing value chain (e.g., equipment manufacturers, operators, and maintenance service providers). Developing adaptive privacy mechanisms that calibrate privacy parameters based on data sensitivity would optimize the privacy-utility tradeoff. Finally, exploring federated transfer learning approaches could address the cold-start problem for new equipment types or facilities joining the federation.

The findings from this research advance both the theoretical understanding of privacy-preserving collaborative intelligence and provide practical implementation frameworks that manufacturing organizations can adopt to enhance predictive maintenance capabilities without compromising sensitive data. This balance between collaboration and competition represents a paradigm shift in how manufacturing intelligence can be developed and deployed across organizational boundaries.

5. Conclusions

This research successfully developed and implemented a federated learning framework for predictive maintenance that addresses the critical challenge of balancing collaborative intelligence with data privacy in manufacturing environments. The framework achieved 93.7% of the predictive performance of centralized approaches while eliminating cross-facility data sharing, demonstrating that privacy-preserving collaborative machine learning is viable in industrial contexts. The implementation across 17 aerospace manufacturing facilities in Uzbekistan resulted in substantial operational improvements, including a 54.2% reduction in production downtime, a 57.2% decrease in emergency maintenance events, and a 22.5% reduction in maintenance costs per machine hour, with a compelling 4.1-month payback period.

By enabling effective predictive maintenance without compromising proprietary process knowledge, this framework creates new opportunities for collaboration across organizational boundaries in manufacturing sectors where intellectual property protection is paramount. The approach bridges the previously insurmountable gap between data isolation and collaborative intelligence, potentially transforming how manufacturing organizations approach maintenance optimization. As manufacturing continues to evolve toward more connected and data-driven paradigms, privacy-preserving approaches like the federated framework presented here will become increasingly critical to balancing competitive advantage with collaborative progress, ultimately advancing the capabilities of smart manufacturing while respecting organizational boundaries.

Acknowledgment

We would like to express my appreciation to everyone involved in supporting this study's completion. We are particularly grateful to the expert advisors who provided critical guidance and perspective throughout the research process.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] G. Lazaroiu, A. Androniceanu, I. Grecu, G. Grecu, and O. Neguriță, “Artificial intelligence-based decision-making algorithms, Internet of Things sensing networks, and sustainable cyber-physical management systems in big data-driven cognitive manufacturing,” Oeconomia Copernicana, vol. 13, no. 4, pp. 1047–1080, 2022, doi: 10.24136/oc.2022.030.

[2] V. Arioli et al., “Digital servitization business typologies in the manufacturing sector,” International Journal of Industrial Engineering and Management, vol. 16, no. 1, pp. 1-23, 2025, doi: 10.24867/IJIEM-378.

[3] A. N. Júnior, P. Nogueira, M. Francescatto, J. Siluk, S. D. Paris, and M. Mandlhate, “Application of a proposed additive manufacturing performance measurement system in a Brazilian industry,” International Journal of Industrial Engineering and Management, vol. 15, no. 2, pp. 109-124, 2024, doi: 10.24867/IJIEM-2024-2-351.

[4] L. Hughes, Y. K. Dwivedi, N. P. Rana, M. D. Williams, and V. Raghavan, “Perspectives on the future of manufacturing within the Industry 4.0 era,” Production Planning & Control, vol. 33, no. 2–3, pp. 138–158, 2022, doi: 10.1080/09537287.2020.1810762.

[5] M. S. Ayubirad, S. Ataei, and M. Tajali, “Numerical Model Updating and Validation of a Truss Railway Bridge considering Train-Track-Bridge Interaction Dynamics,” Shock and Vibration, vol. 2024, no. 1, p. 4469500, 2024, doi: 10.1155/2024/4469500.

[6] M. Siahkouhi, M. Rashidi, F. Mashiri, F. Aslani, and M. S. Ayubirad, “Application of self-sensing concrete sensors for bridge monitoring- A review of recent developments, challenges, and future prospects,” Measurement, vol. 245, p. 116543, 2025, doi: 10.1016/j.measurement.2024.116543.

[7] S. Sajid, A. Haleem, S. Bahl, M. Javaid, T. Goyal, and M. Mittal, “Data science applications for predictive maintenance and materials science in context to Industry 4.0,” Materials today: proceedings, vol. 45, no. 6, pp. 4898–4905, 2021, doi: 10.1016/j.matpr.2021.01.357.

[8] X. Cheng et al., “Systematic literature review on visual analytics of predictive maintenance in the manufacturing industry,” Sensors, vol. 22, no. 17, p. 6321, 2022, doi: 10.3390/s22176321.

[9] M. Alam, M. R. Islam, and S. K. Shil, “AI-Based predictive maintenance for US manufacturing: reducing downtime and increasing productivity,” International Journal of Advanced Engineering Technologies and Innovations, vol. 1, no. 1, pp. 541–567, 2023.

[10] M. Achouch et al., “On predictive maintenance in industry 4.0: Overview, models, and challenges,” Applied Sciences, vol. 12, no. 16, p. 8081, 2022, doi: 10.3390/app12168081.

[11] A. N. Anang and J. N. Chukwunweike, “Leveraging Topological Data Analysis and AI for Advanced Manufacturing: Integrating Machine Learning and Automation for Predictive Maintenance and Process Optimization,” International Journal of Computer Applications Technology and Research, vol. 13, no. 9, pp. 27–39, 2024, doi: 10.7753/IJCATR1309.1003.

[12] A. Bemani and N. Björsell, “Aggregation strategy on federated machine learning algorithm for collaborative predictive maintenance,” Sensors, vol. 22, no. 16, p. 6252, 2022, doi: 10.3390/s22166252.

[13] M. Yazdi, “Maintenance Strategies and Optimization Techniques,” in Advances in Computational Mathematics for Industrial System Reliability and Maintainability, Springer Series in Reliability Engineering. Cham, Switzerland: Springer Nature, 2024, pp. 43–58. doi: 10.1007/978-3-031-53514-7_3.

[14] U. H. W. A. Hewage, R. Sinha, and M. A. Naeem, “Privacy-preserving data (stream) mining techniques and their impact on data mining accuracy: a systematic literature review,” Artificial Intelligence Review, vol. 56, no. 9, pp. 10427–10464, 2023, doi: 10.1007/s10462-023-10425-3.

[15] S. Li, S. Zhao, G. Min, L. Qi, and G. Liu, “Lightweight privacy-preserving scheme using homomorphic encryption in industrial internet of things,” IEEE Internet of Things Journal, vol. 9, no. 16, pp. 14542–14550, 2022, doi: 10.1109/JIOT.2021.3066427.

[16] S. Adelipour and M. Haeri, “Private outsourced model predictive control via secure multi-party computation,” Computers and Electrical Engineering, vol. 116, no. C, p. 109208, 2024, doi: 10.1016/j.compeleceng.2024.109208.

[17] Y. Liu, W. Yu, W. Rahayu and T. Dillon, "An Evaluative Study on IoT Ecosystem for Smart Predictive Maintenance (IoT-SPM) in Manufacturing: Multiview Requirements and Data Quality," IEEE Internet of Things Journal, vol. 10, no. 13, pp. 11160-11184, 2023, doi: 10.1109/JIOT.2023.3246100.

[18] M. S. Azari, F. Flammini, S. Santini and M. Caporuscio, "A Systematic Literature Review on Transfer Learning for Predictive Maintenance in Industry 4.0," IEEE Access, vol. 11, pp. 12887-12910, 2023, doi: 10.1109/ACCESS.2023.3239784.

[19] K. I. -K. Wang, X. Zhou, W. Liang, Z. Yan and J. She, "Federated Transfer Learning Based Cross-Domain Prediction for Smart Manufacturing," IEEE Transactions on Industrial Informatics, vol. 18, no. 6, pp. 4088-4096, 2022, doi: 10.1109/TII.2021.3088057.

[20] R. Mühlhoff, “Predictive privacy: Collective data protection in the context of artificial intelligence and big data,” Big Data & Society, vol. 10, no. 1, 2023, doi: 10.1177/20539517231166886.

[21] N. G. Muminov, R. X. Abdusatarov, A. A. Ambartsumyan, and D. M. Karimov, “Peculiarities of Manufacturing Policy in Uzbekistan in the Conditions of Modernization of the Economy,” Webology, vol. 19, no. 1, pp. 2945–2963, 2022.

[22] N. Nainggolan, E. Maghsoudlou, B. M. AlWadi, F. Atamurotov, M. Kosov, and W. Putra, “Advancements in Optimization for Automotive Manufacturing: Hybrid Approaches and Machine Learning,” International Journal of Industrial Engineering and Management, vol. 15, no. 3, pp. 254-263, 2024, doi: 10.24867/IJIEM-2024-3-361.

[23] D. K. Priatna, W. Roswinna, N. Limakrisna, A. Khalikov, D. Abdullaev, and L. Hussein, “Optimizing Smart Manufacturing Processes and Human Resource Management through Machine Learning Algorithms,” International Journal of Industrial Engineering and Management, vol. 16, no. 2, pp. 176-188, 2025, doi: 10.24867/IJIEM-382.

[24] S. Banabilah, M. Aloqaily, E. Alsayed, N. Malik, and Y. Jararweh, “Federated learning review: Fundamentals, enabling technologies, and future applications,” Information Processing and Management, vol. 59, no. 6, p. 103061, 2022, doi: 10.1016/j.ipm.2022.103061.

[25] P. Qi, D. Chiaro, A. Guzzo, M. Ianni, G. Fortino, and F. Piccialli, “Model aggregation techniques in federated learning: A comprehensive survey,” Future Generation Computer Systems, vol. 150, pp. 272–293, 2024, doi: 10.1016/j.future.2023.09.008.

[26] B. Alotaibi, F. A. Khan, and S. Mahmood, “Communication Efficiency and Non-Independent and Identically Distributed Data Challenge in Federated Learning: A Systematic

Mapping Study," Applied Sciences, vol. 14, no. 7, p. 2720, 2024, doi: 10.3390/app14072720.

[27] H. Zhu, H. Zhang, and Y. Jin, "From federated learning to federated neural architecture search: a survey," Complex & Intelligent Systems, vol. 7, no. 2, pp. 639–657, 2021, doi: 10.1007/s40747-020-00247-z.

[28] A. Blika et al., "Federated Learning for Enhanced Cybersecurity and Trustworthiness in 5G and 6G Networks: A Comprehensive Survey," IEEE Open Journal of the Communications Society, vol. 6, pp. 3094-3130, 2025, doi: 10.1109/OJCOMS.2024.3449563

[29] X. Yuan et al., "FedComm: A Privacy-Enhanced and Efficient Authentication Protocol for Federated Learning in Vehicular Ad-Hoc Networks," IEEE Transactions on Information Forensics and Security, vol. 19, pp. 777-792, 2024, doi: 10.1109/TIFS.2023.3324747.

[30] R. Shokri, M. Stronati, C. Song and V. Shmatikov, "Membership Inference Attacks Against Machine Learning Models," 2017 IEEE Symposium on Security and Privacy (SP), San Jose, CA, USA, 2017, pp. 3-18, doi: 10.1109/SP.2017.41.

[31] A. Chaddad, Y. Wu and C. Desrosiers, "Federated Learning for Healthcare Applications," IEEE Internet of Things Journal, vol. 11, no. 5, pp. 7339-7358, 2024, doi: 10.1109/JIOT.2023.3325822.

[32] T. Deng, Y. Li, X. Liu, and L. Wang, "Federated learning-based collaborative manufacturing for complex parts," Journal of Intelligent Manufacturing, vol. 34, no. 7, pp. 3025–3038, 2023, doi: 10.1007/s10845-022-01968-3.

[33] C. Wu, F. Wu, L. Lyu, Y. Huang, and X. Xie, "Communication-efficient federated learning via knowledge distillation," Nature Communications, vol. 13, no. 1, p. 2032, 2022.

[34] G. K. Jagarlamudi, A. Yazdinejad, R. M. Parizi, and S. Pouriyeh, "Exploring privacy measurement in federated learning," Journal of Supercomputing, vol. 80, no. 8, pp. 10511–10551, 2024, doi: 10.1007/s11227-023-05846-4.