Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 05 , May - 2026

Federated Learning-Based Intelligent Product Lifecycle Management Framework for Automotive Supply Chains: Enhancing Data Privacy, Predictive Analytics and Collaborative Decision-Making

Paladi Gopikrishna
Research Scholar
ISBR Business School, Bangalore, India

***Abstract* - The automotive industry is experiencing a major transformation driven by digital technologies, distributed manufacturing systems and strict data governance requirements. Traditional Product Lifecycle Management (PLM) systems face limitations due to centralized data structures that hinder collaboration across organizations. They also pose significant risks to intellectual property and competitive insights. This study introduces a new Federated Learning-Based Intelligent PLM (FL-iPLM) framework designed specifically for automotive supply chains. It allows various stakeholders, including original equipment manufacturers (OEMs), tier-1 suppliers, logistics partners and aftermarket service providers, to collaboratively develop shared predictive models without exposing sensitive data. The framework combines federated aggregation protocols, privacy-preserving techniques and a multilayer digital twin architecture to improve predictive maintenance, demand forecasting and quality assurance. A hierarchical aggregation strategy with adaptive client weighting addresses the issue of statistical differences in data across supply chain nodes. Experiments carried out within a simulated automotive supply chain comprising six organizational entities demonstrated that the FL-iPLM framework attained predictive accuracy levels that differed by only 2.1% from those of centralized baseline models. It also reduced data exposure risk by 94.7% and improved decision-making speed between organizations by 38.4%. The framework supports the principles of Industry 5.0, focusing on human-centered collaboration and sustainable manufacturing. The results indicate that this approach provides a practical, scalable and privacy-compliant pathway for the future of automotive PLM.**

Keywords - federated learning; product lifecycle management; automotive supply chain; differential privacy; Industry 5.0; predictive analytics; digital twin; non-IID data

I. INTRODUCTION

The global automotive sector is one of the most complex and widely dispersed manufacturing systems. Producing a single vehicle typically involves over 30,000 individual parts sourced from hundreds of suppliers across several continents. These parts are coordinated through intricate Product Lifecycle Management (PLM) systems that oversee design, engineering, production, quality and aftermarket operations [1]. The digital transformation of these systems has been sped up by Industry 4.0 requirements and further shaped by the human-focused goals of Industry 5.0. This shift demands that PLM platforms change from isolated, OEM-centered systems into intelligent, collaborative and privacy-sensitive decision-support environments [2].

Centralized machine learning (ML) approaches show much potential in improving PLM functions like predictive maintenance, defect classification, demand forecasting and supply chain risk assessment. However, the use of centralized AI in automotive supply chains faces three main challenges. First, concerns about data ownership prevent OEMs, tier-1 suppliers and logistics partners from sharing raw operational data with centralized platforms controlled by competitors or third-party vendors. Second, regulations like the General Data Protection Regulation (GDPR), China's Personal Information Protection Law (PIPL) and industry-specific cyber security standards impose strict limits on cross-border data transfers [3]. Finally, the costs of gathering multi-source, high-frequency operational data at centralized servers become prohibitive at larger scales.

Federated Learning (FL), first introduced by McMahan et al. [4], addresses the issue of centralized data aggregation. It enables collaborative model training among distributed clients, where only model updates, rather than raw data, are sent to an aggregation server. This approach has been explored in healthcare, finance and IoT applications, yet its application to automotive PLM's distinct challenges—including varying data distributions across supply chain tiers, differing client participation and the need for quick decision-making—remains underdeveloped in the research.

This study contributes to the field in four main ways. First, it proposes FL-iPLM, a federated learning-based PLM framework specifically created for multi-tier automotive supply chains. This framework includes a hierarchical aggregation structure that reflects the OEM-supplier-logistics hierarchy. Second, it introduces an Adaptive Federated Aggregation (AFA) algorithm that addresses the issues caused by non-independent and identically distributed (non-IID) data. This is done through weighting client contributions based on local data quality. Third, it incorporates a multilayer digital twin module within the federated framework. This enables real-time synchronization of physical asset states across organizational boundaries without compromising data privacy. Fourth, it presents a thorough experimental evaluation in a simulated six-node automotive supply chain environment, showcasing the framework's effectiveness in predictive maintenance, quality assurance and demand forecasting.

The rest of this paper is organized as follows: Section II reviews related work on federated learning and automotive PLM. Section III describes the proposed FL-iPLM framework. Section IV explains the Adaptive Federated Aggregation algorithm. Section V outlines the experimental methods and results. Section VI discusses the implications for Industry 5.0 adoption. Section VII concludes with future research directions.

II. RELATED WORK

A. Product Lifecycle Management in Automotive Contexts

PLM in automotive manufacturing has evolved from engineering data management (EDM) systems into comprehensive digital systems that include CAD/CAM integration, bill-of-materials management, configuration control and aftermarket service analytics [5]. Nowadays, automotive PLM platforms like Siemens Teamcenter, PTC Windchill and Dassault ENOVIA are increasingly integrating AI-driven features. However, these systems remain mainly centralized and focused on OEMs, creating significant obstacles to sharing data across organizations [6].

Several researchers have pointed out the limitations of current PLM structures in handling the distributed nature of modern automotive supply chains. Kiritsis [7] argued for a shift towards closed-loop product lifecycle management, emphasizing the disconnect between product data generated across supply chain tiers and the centralized PLM systems maintained by OEMs. Subramaniam et al. [8] proposed a multi-agent framework for distributed PLM but did not address the privacy issues that arise with real-world cross-organizational implementations.

B. Federated Learning: Foundations and Applications

The foundational federated averaging (FedAvg) algorithm introduced by McMahan et al. showed that efficient distributed ML could be achieved by collecting locally trained model updates at a central server. Later work refined this method to tackle limitations such as communication overhead, client drift and convergence issues with non-IID data distributions. Li et al. [9] proposed FedProx, which adds a constraint to local model updates, improving convergence in varied settings—a key factor for automotive supply chain environments where data distributions can differ greatly.

Privacy amplification techniques have been incorporated into federated systems through differential privacy mechanisms [10], secure multiparty computation [11] and homomorphic encryption [12]. Geyer et al. [13] demonstrated client-level privacy within federated learning, proving that meaningful privacy guarantees can be reached alongside acceptable accuracy. These privacy-enhancing methods are particularly relevant for automotive supply chains, where even aggregated gradient information can reveal sensitive production volumes or process parameters.

C. Federated Learning in Manufacturing and Supply Chain Contexts

The use of FL in industrial manufacturing has gained traction in recent years. Liu et al. [14] applied federated learning to predictive maintenance across multiple manufacturing sites, showing accuracy improvements over models tailored for specific locations because of the broader training signal provided by shared data. Zhang et al. [15]. examined FL-based quality inspections in semiconductor manufacturing, achieving defect detection rates similar to centralized methods while maintaining data confidentiality across different plants.

In supply chain management, Cheng et al. [16] introduced a federated demand forecasting system for retail supply chains using long short-term memory (LSTM) networks. Their results indicated that federated models outperformed individually trained local ones, particularly for nodes with limited historical data. However, there is still a lack of studies specifically addressing the needs of the automotive PLM domain, including engineering change management, multi-tier supplier collaboration and tracking of components over time. This is a primary reason for this study.

III. PROPOSED FL-IPLM FRAMEWORK

A. Framework Overview

The FL-iPLM framework is designed as a three-tier hierarchical structure that corresponds with the organizational layout of a typical automotive supply chain: (i) the Edge Tier, consisting of individual supplier nodes and production facilities that host locally trained models and Digital Twin clients; (ii) the Cluster Aggregation Tier, made up of intermediate servers located with tier-1 suppliers or regional logistics hubs; and (iii) the Global Coordination Tier, managed by a neutral trusted coordinator—like an industry consortium or OEM federation—responsible for global model aggregation and policy governance.

This hierarchical structure has two main advantages over flat federated systems. First, it cuts down on communication delays and bandwidth use by performing aggregation at nearby nodes before sending consolidated updates to the global coordinator. Second, it allows for independent management of privacy budgets at the cluster level. This enables suppliers within competitive groups to share models without exposing updates to rival OEM supply chains.

B. Digital Twin Integration

A key element for real-time decision-making in PLM is the use of Digital Twin (DT) technology within a federated framework. In FL-iPLM, each supply chain node has a Digital Twin client that consistently syncs physical asset states, such as machine health indices, component tracing, process parameters and quality inspection records, from shop-floor sensors and enterprise resource planning (ERP) systems. The DT layer accomplishes three roles in the federated architecture: (i) local data preprocessing and feature engineering before local model training; (ii) simulation of scenarios to enhance sparse training datasets through synthetic data creation; and (iii) real-time anomaly detection using locally trained models to prompt maintenance or quality checks without needing global model evaluations.

C. Privacy-Preserving Mechanisms

FL-iPLM features a layered privacy design that uses three complementary methods. At the local training level, Gaussian differential privacy noise is added to gradient updates before they are sent, with the privacy budget ε adjusted for each communication round using a privacy accountant based on the Renyi differential privacy (RDP) framework [17]. At the cluster aggregation level, secure aggregation through masked model updates prevents any single aggregator node from seeing individual client contributions in plaintext. At the global coordination level, an audit trail based on cryptographic commitments lets regulators or consortium members check the integrity of the aggregation without accessing the underlying model parameters.

The privacy budget allocation strategy in FL-iPLM is flexible. Nodes with more sensitive data categories, such as proprietary tooling parameters or contractual production volumes, receive smaller ε values for stronger privacy. In contrast, nodes that provide less sensitive operational telemetry can use larger ε budgets, which helps enhance their contribution to model convergence. This tiered privacy allocation follows a data sensitivity classification system defined in the consortium's PLM governance policy.

D. Framework System Architecture

Tier	Node Type	Primary Function	Privacy Mechanism
Edge	Supplier / Plant	Local model training, DT synchronization	Differential Privacy (Gaussian noise)
Cluster	Tier-1 Aggregator	Intra-cluster aggregation, anomaly escalation	Secure Aggregation (masking)
Global	Consortium Coordinator	Global FedAvg, policy governance, audit	Cryptographic commitments

TABLE I. FL-iPLM Framework Tier Specifications

IV. ADAPTIVE FEDERATED AGGREGATION ALGORITHM

A. Motivation: Addressing Non-IID Data Heterogeneity

A key challenge in applying federated learning to automotive supply chains is the significant variation in data across nodes. An engine component supplier mainly generates time-series data on vibrations and temperatures from machining operations. In contrast, a logistics partner contributes delivery telemetry, transit condition logs and throughput metrics. An OEM quality inspection node provides image-based defect classification datasets and dimension measurement records. Standard FedAvg, which combines model updates weighted by local dataset size, can show considerable client drift under such non-IID distributions, leading to slow convergence or poorer quality in the global model [9].

B. AFA Algorithm Formulation

The proposed Adaptive Federated Aggregation (AFA) algorithm builds on FedAvg by adding a client weighting function that considers three factors: data quantity (nk), data quality score (qk) and model gradient alignment (αk). The global model update at round t is calculated as follows:

$$w^{(t+1)} = \Sigma_k \left[(n_k \cdot q_k \cdot \alpha_k) / (\Sigma_j n_j \cdot q_j \cdot \alpha_j) \right] \cdot w_k^{(t)}$$

Here, $w_k^{(t)}$ represents the local model weights from client k after the local training round t. The values nk and qk ∈ [0,1] denote the number of local training samples and a data quality metric based on completeness, consistency and recency scores. The value αk ∈ [-1,1] indicates the cosine similarity between the local gradient update and the global gradient direction from the previous round, serving as a proxy for gradient alignment that penalizes updates diverging from the global objective.

The gradient alignment score αk has two purposes: detecting Byzantine clients (nodes that submit manipulated gradient updates) and naturally down-weighting clients whose local data distributions temporarily misalign with the global distribution due to operational issues, such as a plant undergoing retooling or producing an unusually high number of defective units. This built-in robustness mechanism is particularly useful in automotive supply chains, where nodes might show atypical data distributions without being involved in adversarial behavior.

C. Communication Efficiency Optimization

To minimize communication costs, a critical consideration for automotive suppliers operating across heterogeneous network environments, FL-iPLM incorporates a gradient compression mechanism. This approach integrates top-k sparsification, retaining only the most significant k gradient values in each layer, alongside 8-bit fixed-point quantization to reduce data transmission requirements. This approach achieves compression ratios of 15–25× compared to full-precision gradient transmission, allowing suppliers with limited bandwidth to participate while maintaining convergence properties with an acceptable accuracy penalty of less than 0.8% across evaluated tasks.

FL-iPLM also supports asynchronous participation through a staleness-tolerant aggregation buffer. This buffer accepts gradient updates from clients for up to τ rounds, weighted by a staleness decay factor δ^(t - tk), where tk is the round when client k last submitted an update. This method accounts for production shifts, maintenance windows, or planned downtimes that may disrupt regular FL client participation.

V. EXPERIMENTAL EVALUATION

A. Experimental Setup

To assess the FL-iPLM framework, a simulated automotive supply chain environment was created with six organizational nodes: one OEM assembly plant (Node O1), two tier-1 suppliers (Nodes S1: power train components, S2: chassis and body systems), one tier-2 supplier (Node S3: electronic control units), one logistics and distribution partner (Node L1) and one aftermarket service network aggregator (Node A1). Each node operated as an independent federated learning client with isolated data partitions relevant to its operational domain.

Three PLM-related prediction tasks were conducted: (i) Predictive Maintenance (PM) — binary classification of potential machine failures within a 72-hour window, using the NASA CMAPSS turbofan degradation dataset adjusted to simulate automotive machining equipment; (ii) Demand Forecasting (DF) — multi-step time-series forecasting of component demands across the supply chain with a synthetic dataset based on automotive production planning patterns; and (iii) Quality Defect Detection (QD) — multi-class classification of surface and dimension defects in machined components using simulated coordinate measuring machine (CMM) and optical inspection data.

B. Baseline Comparisons

FL-iPLM was compared against four baseline setups: (1) Local-only models trained independently on each node's data without federation; (2) Centralized models trained on the combined dataset from all nodes, representing the best performance without privacy constraints; (3) Standard FedAvg [4]; and (4) FedProx [9] with a proximal regularization coefficient $\mu = 0.01$. All federated configurations were assessed over 100 global communication rounds with five local epochs per round, using the Adam optimizer with an initial learning rate of 0.001.

C. Results and Analysis

Table II shows the predictive performances of FL-iPLM and the baseline methods across the three evaluation tasks. FL-iPLM achieved F1 scores of 0.923, 0.891 and 0.912 for the PM, DF and QD tasks respectively. This compares to the centralized baseline scores of 0.941, 0.908 and 0.931. The average accuracy gap of 2.1% compared to the centralized baseline indicates a substantial improvement over standard FedAvg (4.8% gap) and FedProx (3.3% gap). This confirms the effectiveness of the AFA algorithm in reducing the impact of non-IID distributions.

Method	**PM (F1)**	**DF (MAE↓)**	**QD (F1)**	**Avg. Gap vs. Centralized**
Local-Only	0.804	18.72	0.813	12.4%
Centralized	0.941	11.34	0.931	—
FedAvg [4]	0.887	14.81	0.896	4.8%
FedProx [9]	0.903	13.52	0.912	3.3%
FL-iPLM (Proposed)	0.923	12.19	0.912	2.1%

TABLE II. Predictive Performance Comparison Across Evaluation Tasks

The privacy analysis shows that FL-iPLM, with differential privacy ($\varepsilon = 2.0$, $\delta = 10^{-5}$), lowers data exposure risk. This is shown as the mutual information between sent gradient updates and raw training data, reduced by 94.7% compared to centralized training and by 71.3% against standard FedAvg without explicit DP methods. This reduction occurred with a small accuracy drop of 1.2% relative to FL-iPLM without DP, indicating that the privacy-accuracy balance is manageable within the framework.

Latency for cross-organization decisions, measured as the time from detecting an anomaly at an edge node to sending coordinated response recommendations to all affected supply chain partners, decreased by 38.4% in FL-iPLM compared to the baseline sequential communication method used in traditional centralized PLM systems. This improvement comes from localized DT-based anomaly detection, allowing nodes to start preliminary responses without waiting for global model inference from a centralized server.

D. Scalability Analysis

We evaluated the scalability of FL-iPLM by gradually increasing the number of federated clients from 6 to 50, simulating a transition from a pilot program to a complete tier-1 supplier network. The communication burden per global round scaled sub-linearly with the number of clients, increasing by a factor of 3.2× for an 8.3× increase in client count. Model convergence needed an average of 18 additional communication rounds to achieve similar accuracy when expanding from 6 to 50 clients. This modest overhead aligns with the theoretical convergence limits of FedAvg-based algorithms under partial client participation.

VI. INDUSTRY 5.0 IMPLICATIONS AND DEPLOYMENT CONSIDERATIONS

A. Alignment with Industry 5.0 Principles

Industry 5.0 builds on the automation-focused vision of Industry 4.0 by putting human expertise and ethical responsibility back at the center of digital manufacturing systems [18]. The FL-iPLM framework aligns explicitly with three main Industry 5.0 pillars: (i) human-centricity, where federated model outputs are interpreted through explain ability layers using SHapley Additive exPlanations (SHAP) attributions. This allows domain engineers to validate and understand AI recommendations before making operational decisions; (ii) sustainability, where the hierarchical aggregation structure minimizes unnecessary data transmission and energy use compared to centralized methods; and (iii) resilience, where the distributed federated design removes single points of failure that could disrupt supply chain nodes during server outages or cyber attacks.

B. Governance and Consortium Requirements

To deploy FL-iPLM successfully, a formal Industry Data Consortium (IDC) must be established. This consortium will be governed by a multilateral agreement that outlines model architecture standards, communication protocols, privacy budget allocations, audit rights and intellectual property ownership of the global model. The automotive industry has examples of such consortia, like the Catena-X data ecosystem, which offers a governance framework that FL-iPLM could integrate with through standardized API interfaces [19]

C. Integration with Legacy PLM Systems

A major deployment challenge is integrating FL-iPLM with existing PLM platforms that have been in use for decades and are deeply woven into organizational workflows. The proposed framework addresses this by using an abstraction layer that provides standardized REST APIs compatible with major PLM platforms. This lets FL client modules access data and return recommendations to current PLM repositories without needing full system replacements. This integration-first approach lowers deployment risks and enables step-by-step adoption, starting with high-priority PLM functions like predictive maintenance, then expanding to more complex collaborative decision-making scenarios.

VII. CONCLUSION

This study introduced FL-iPLM, a federated learning-based intelligent Product Lifecycle Management framework designed for the privacy-sensitive and complex environment of automotive supply chains. Its hierarchical federated architecture, Adaptive Federated Aggregation algorithm, integrated digital twin module and layered privacy measures address the main challenges facing AI adoption in cross-organizational automotive PLM: concerns over data sovereignty, statistical differences across supply chain nodes and communication infrastructure issues.

Experimental results across three critical PLM prediction tasks show that FL-iPLM achieves predictive accuracy within 2.1% of an ideal centralized baseline. It also provides a 94.7% reduction in data exposure risk and a 38.4% improvement in decision-making speed across organizations. These findings make a strong case for adopting federated learning as the key technology for next-generation collaborative automotive PLM. This is especially relevant within the evolving Industry 5.0 framework, which demands enhanced AI capability and increased human and societal accountability.

Future work will focus on three areas: (i) creating personalized federated learning variants that let individual supply chain nodes keep their specific model adaptations alongside the global federated model, improving local prediction accuracy for nodes with highly specific data distributions; (ii) integrating FL-iPLM with blockchain-based smart contracts to automate trustless privacy budget governance and create incentives for high-quality client contributions; and (iii) validating the framework in partnership with an industrial automotive consortium, moving from simulations to real-world use within a pilot supplier network.

ACKNOWLEDGMENT

The author appreciates the support of the ISBR Business School in Bengaluru. The author also thanks the anonymous reviewers for their helpful feedback, which significantly enhanced the manuscript's quality. No financial assistance or external funding was provided for the completion of this study.

REFERENCES

[1] Grieves, M. (2005). Product lifecycle management: The new paradigm for enterprises. International Journal of Product Development, 2(1–2), 71–84.

[2] European Commission, "Industry 5.0: Towards a sustainable, human-centric and resilient European industry," Publications Office of the European Union, Luxembourg, 2021.

[3] A.Chander, "The crypto-currency risk," Columbia Law Review, vol. 111, no. 3, pp. 1–30, 2011; cf. R. Voigt et al., "The GDPR: Requirements, choices and expected impact," Business Information Systems Engineering, vol. 62, no. 4, pp. 279–287, 2020.

[4] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 2017, pp. 1273–1282.

[5] J. Stark, Product Lifecycle Management: 21st Century Paradigm for Product Realisation, 3rd ed. Geneva, Switzerland: Springer, 2015.

[6] M. Alemanni, F. Gracis and E. Tornincasa, "A new paradigm for the definition of smart product concepts in the PLM framework," in Proc. 7th Int. Conf. PLM, Brisbane, Australia, 2010, pp. 200–209.

[7] D.Kiritsis, "Closed-loop PLM for intelligent products in the era of the Internet of Things," Computer-Aided Design, vol. 43, no. 5, pp. 479–501, 2011.

[8] A. Subramaniam, P. Venkatesan and K. Ramachandran, "Multi-agent systems for distributed product lifecycle management: An architectural perspective," Computers in Industry, vol. 131, p. 103482, 2021.

[9] Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). *Federated optimization in heterogeneous networks*. Presented at the Machine Learning and Systems Conference (MLSys), held in Austin, Texas, USA.

[10] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," Foundations and Trends in Theoretical Computer Science, vol. 9, nos. 3–4, pp. 211–407, 2014.

[11] P. Mohassel and Y. Zhang, “SecureML: A system for scalable privacy-preserving machine learning,” in *Proceedings of the IEEE Symposium on Security and Privacy (S&P)*, 2017, pp. 19–38.

[12] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal and K. Seth, "Practical secure aggregation for privacy-preserving machine learning," in Proc. ACM SIGSAC Conf. Comput. Commun. Security (CCS), Dallas, TX, USA, 2017, pp. 1175–1191.

[13] R. C. Geyer, T. Klein and M. Nabi, "Differentially private federated learning," arXiv preprint arXiv:1712.07557, 2017.

[14] Liu, Y., Chen, T., & Yang, Q. (2020). Secure federated transfer learning. *IEEE Intelligent Systems, 35*(4), 70–82.

[15] W. Zhang, X. Li, R. Ma, Z. Zhan and X. He, "Federated learning for machinery fault diagnosis with dynamic validation and self-supervision," Knowledge-Based Systems, vol. 213, p. 106679, 2021.

[16] Y. Cheng, Y. Lu, K. Wang, and H. Liang, “Federated learning for demand forecasting in heterogeneous retail supply chains,” *International Journal of Production Economics*, vol. 248, p. 108493, 2022.

[17] Mironov, I. (2017). *Rényi differential privacy*. In *Proceedings of the 30th IEEE Computer Security Foundations Symposium (CSF)* (pp. 263–275).

[18] R. Breque, L. De Nul and A. Petridis, "Industry 5.0: Towards a sustainable, human-centric and resilient European industry," European Commission, Directorate-General for Research and Innovation, Brussels, 2021.

[19] Catena-X Automotive Network, "Catena-X — The automotive data ecosystem: Collaborative standards for data spaces," Catena-X Automotive Network e.V., Berlin, Germany, 2023. [Online]. Available: https://catena-x.net