International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 15 Issue 3, March-2026

A Privacy-Aware Federated Learning Model for Secure Data Collaboration in Smart Living Portal

Ms.Anushka Prasad Joshi[1], Prof.Sachin Bhosale[2], Dr.Anand Khatri[3],Ms.Pallavi Bhatt[4]

[1]PG Research Scholar, dept. of Computer Engineering ,Jaihind College OfEngineering , Pune University,Pune,Maharashtra,India
[2,3]Professor, dept. of Computer Engineering ,Jaihind College OfEngineering , Pune University,Pune, Maharashtra,India
[4]Professor,dept. of Computer Science, BanarsidasChandiwala Institute of Information Technology, Guru Gobind Singh Indraprastha University, New Delhi,India

ABSTRACT

The increasing adoption of smart living technologies has resulted in continuous data generation from connected devices such as sensors, appliances, and home automation systems. Although centralized data processing supports intelligent services and predictive analytics, it also raises significant concerns related to user privacy, data misuse, and unauthorized access. To overcome these challenges, this study proposes a privacy-aware federated learning model tailored for secure data collaboration in smart living portals.

The proposed framework enables multiple smart devices and local systems to train a shared machine learning model without transmitting raw data to a central server. Instead, each participant performs local model training and shares only model parameters or updates through protected communication channels. Additional privacy safeguards, including secure aggregation and controlled noise injection techniques, are incorporated to minimize the risk of sensitive information exposure. The architecture is designed to operate efficiently in heterogeneous smart environments and supports scalability across multiple residential units. Experimental evaluation demonstrates that the framework achieves competitive model performance while maintaining strong privacy protection. The results indicate that distributed learning can significantly reduce data vulnerability without compromising analytical accuracy. The proposed model offers a practical and secure solution for intelligent data integration in smart habitation ecosystems, contributing toward the development of trustworthy and privacy-conscious smart living platforms.

Keywords: **Federated Learning, Privacy Preservation, Smart Living Systems, Secure Data Collaboration, Distributed Machine Learning, Smart Home Ecosystems, Secure Aggregation, Data Protection.**

INTRODUCTION

The concept of smart living has rapidly evolved with the advancement of Internet of Things (IoT) technologies, cloud computing, and intelligent data analytics. Modern smart habitation ecosystems integrate connected devices such as environmental sensors, surveillance systems, wearable health monitors, and automated appliances to enhance comfort, energy efficiency, safety, and overall quality of life. These devices continuously generate large volumes of data that are processed through smart living portals to provide personalized and automated services.

Traditionally, such data are collected and transmitted to centralized servers for storage and analysis. While centralized machine learning models can extract valuable insights, this approach exposes sensitive personal information to potential privacy breaches, unauthorized access, and cyber threats. Smart home data often include behavioural patterns, health-related information, energy consumption habits, and security logs, making privacy protection a critical requirement. Moreover, increasing regulatory frameworks and user awareness regarding data protection demand solutions that minimize direct data sharing.

Federated Learning (FL) has emerged as a promising distributed machine learning paradigm that addresses these privacy concerns. Instead of transferring raw data to a central location, federated learning enables devices to train models locally and share only model updates with a coordinating server. This decentralized approach reduces the risk of data leakage while still enabling collaborative intelligence. However, despite its advantages, federated learning faces challenges related to secure communication, model update leakage, system heterogeneity, scalability, and computational constraints within smart environments.

In the context of smart living portals, the integration of privacy-aware federated learning requires additional mechanisms such as secure aggregation protocols, encryption techniques, and noise-based privacy preservation strategies to strengthen data confidentiality. Ensuring reliable performance while maintaining privacy guarantees is essential for practical deployment in real-world residential ecosystems.

This research proposes a privacy-aware federated learning framework specifically designed for secure data collaboration in smart habitation ecosystems. The framework aims to enable intelligent decision-making without compromising user confidentiality. It focuses on building a scalable and secure architecture that balances model accuracy, communication efficiency, and privacy protection.

RELATED WORK

The rapid expansion of smart living environments has encouraged extensive research in privacy-preserving data analytics and distributed machine learning. Traditional smart home systems rely on centralized cloud architectures for storing and processing user data. Although such systems offer high computational power and simplified model management, several studies have highlighted the associated risks of data breaches, unauthorized access, and privacy violations in centralized frameworks. These concerns have motivated researchers to explore decentralized and privacy-aware learning approaches. Federated Learning (FL) was introduced as a distributed machine learning paradigm that allows multiple clients to collaboratively train a global model without sharing raw data. Early research in federated learning demonstrated its effectiveness in mobile devices and edge computing environments, where local data remain on-device while only model updates are transmitted to a central aggregator. This approach significantly reduces direct exposure of personal information and supports compliance with privacy regulations. However, subsequent studies revealed that model updates themselves may leak sensitive information through inference attacks.

To strengthen privacy guarantees, researchers incorporated techniques such as differential privacy and secure multi-party computation into federated learning frameworks. Differential privacy introduces controlled statistical noise to model updates, limiting the ability of adversaries to extract individual data points. Secure aggregation protocols were also proposed to ensure that the central server can only access aggregated model parameters rather than individual client updates. These enhancements improved confidentiality but often increased computational overhead and communication cost.

In smart home and smart city ecosystems, federated learning has been explored for applications such as energy consumption prediction, health monitoring, anomaly detection, and intelligent automation. Several works have focused on integrating edge computing with federated learning to reduce latency and bandwidth usage. Despite these advancements, challenges remain in handling heterogeneous devices, unstable network connections, and varying computational capacities within smart habitation environments.

Recent research efforts emphasize building lightweight and scalable federated architectures tailored for IoT-based ecosystems. Hybrid models combining encryption techniques, blockchain-based integrity verification, and adaptive model aggregation strategies have been proposed to enhance trust and security. Nevertheless, achieving an optimal balance between privacy preservation, model accuracy, system efficiency, and real-time responsiveness continues to be an open research problem.

Based on the existing literature, it is evident that while federated learning provides a strong foundation for privacy-preserving collaboration, there is a need for a specialized framework designed specifically for smart living portals. The proposed work addresses this gap by integrating privacy-aware mechanisms with a scalable federated learning architecture suitable for heterogeneous smart habitation ecosystems.

METHODOLOGY

This section describes the proposed privacy-aware federated learning framework designed for secure data collaboration in smart living portals. The methodology focuses on distributed model training, secure communication, and privacy enhancement while maintaining system efficiency.

3.1 System Architecture Overview

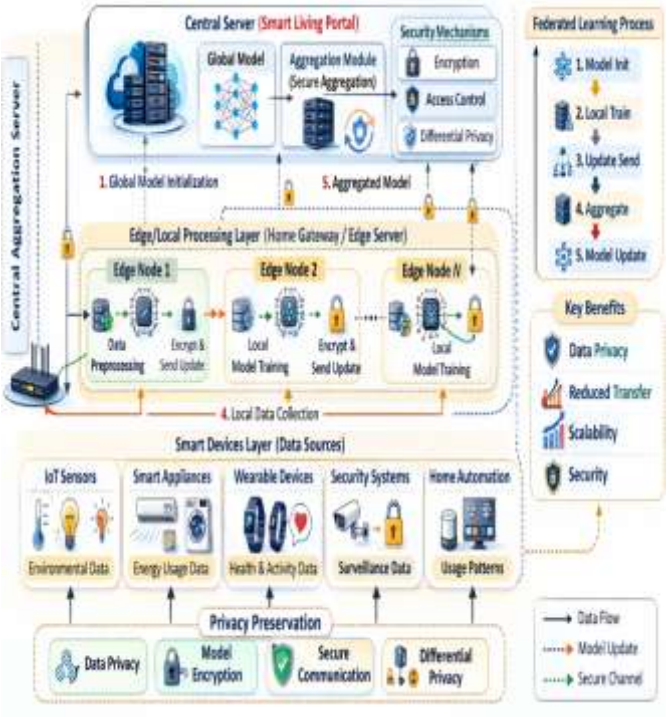

Fig.1:proposed privacy-aware federated learning framework designed for secure data collaboration

The proposed framework consists of three major components:

1. **Smart Devices Layer** – Includes IoT sensors, wearable devices, smart appliances, and monitoring systems deployed in residential environments. These devices generate local data such as energy usage, environmental conditions, occupancy patterns, and health indicators.
2. **Edge/Local Processing Layer** – Each smart home unit contains a local processing node (edge server or gateway) responsible for:
   - Preprocessing raw data
   - Training a local machine learning model
   - Encrypting model updates before transmission
3. **Central Aggregation Server (Smart Living Portal)** – This server coordinates the federated learning process by:
   - Initializing the global model
   - Collecting encrypted local model updates
   - Performing secure aggregation
   - Updating and redistributing the global model

Raw user data never leave the local environment, ensuring privacy protection.

3.2 Data Preprocessing

Before training, each local node performs preprocessing steps:

- Data cleaning (removal of missing or inconsistent values)
- Feature normalization and scaling
- Feature selection to reduce computational load
- Data labelling (if supervised learning is applied)

This step ensures consistent model performance across heterogeneous devices.

3.3 Federated Learning Process

The federated learning process is executed iteratively in communication rounds:

Step1:GlobalModel Initialization

The central server initializes a base machine learning model and distributes it to participating local nodes.

Step 2: Local Model Training

Each smart node trains the model using its own local dataset for a predefined number of epochs.

Step 3: Privacy Protection Mechanism

Before sending updates:

- Model parameters are encrypted.
- Controlled noise is added using differential privacy techniques.
- Secure aggregation protocols are applied to prevent exposure of individual updates.

Step 4: Secure Aggregation

The central server aggregates the received updates using a weighted averaging algorithm to generate an improved global model.

Step 5: Global Model Update

The updated model is redistributed to all participants for the next training round.

This process continues until convergence criteria (such as accuracy threshold or maximum rounds) are met.

3.4 Privacy-Preserving Mechanisms

To enhance confidentiality, the framework integrates:

- **Secure Aggregation:** Ensures the server cannot access individual client updates.
- **Differential Privacy:** Adds statistical noise to limit information leakage.
- **Encrypted Communication Channels:** Protects data transmission from interception.
- **Access Control Policies:** Restrict unauthorized participation in the training process.

These mechanisms collectively reduce risks of inference attacks and model inversion attacks.

3.5 Performance Evaluation Metrics

The framework is evaluated using:

- Model Accuracy
- Precision, Recall, and F1-Score
- Communication Overhead
- Convergence Rate
- Privacy Leakage Risk Indicators

This multi-metric evaluation ensures that privacy improvements do not significantly degrade learning performance.

3.6 Scalability and Adaptability

The methodology supports:

- Heterogeneous device participation
- Dynamic joining and leaving of nodes
- Resource-aware model updates
- Low-bandwidth environments

This makes the framework suitable for real-world smart habitation ecosystems.

RESULTS

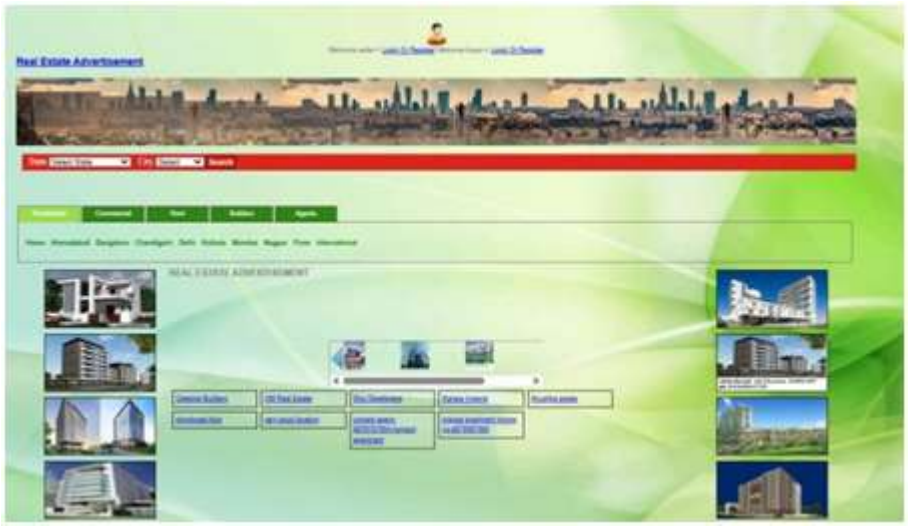

Fig.2:SMART Habitation-Real estate Advertising Advertisement Portal

The Real Estate Advertisement Portal diagram illustrates the structural design and operational workflow of a web-based property management and advertisement system within a smart habitation ecosystem. The portal acts as a centralized user interface that enables buyers, sellers, and agents to interact with real estate listings in an organized and efficient manner. The system provides search and filtering mechanisms based on parameters such as state, city, budget, and property category, including residential, commercial, rental, and international listings. The interface is designed to enhance user accessibility and streamline property discovery through categorized navigation and structured display sections. The portal integrates dynamic advertisement modules that showcase featured properties along with images and key specifications. Behind the user interface, the system connects to a property database and a recommendation engine that may incorporate federated learning mechanisms for personalized suggestions. This ensures that users receive relevant property options based on their preferences and search behaviour while maintaining data privacy. Overall, the architecture supports scalability, user engagement, and intelligent property management through structured information flow between users, the web interface, and backend processing systems.

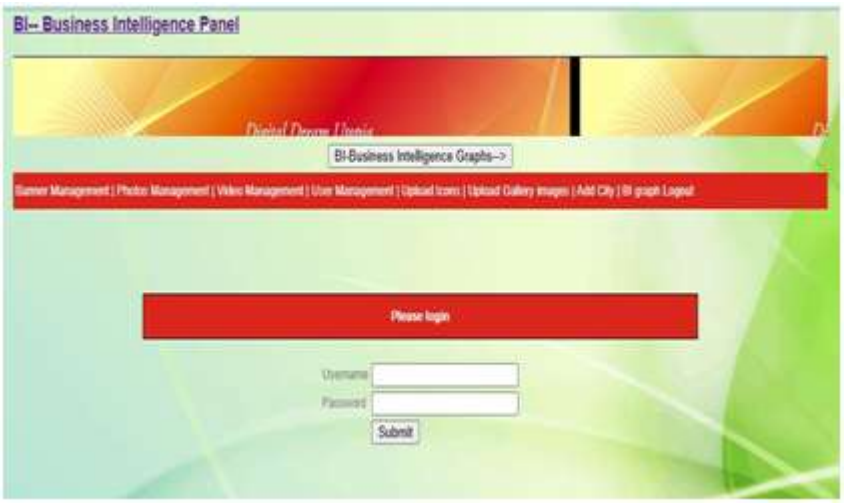

Fig.3: UseOf BI Panel for Performance evaluation

The Business Intelligence (BI) Panel diagram represents the analytical and administrative component of the real estate system. This module is designed to provide data-driven insights derived from aggregated and processed property and user interaction data. Access to the BI dashboard is secured through an authentication mechanism, ensuring that only authorized administrators can manage system operations and analyze performance metrics.

The BI panel includes management functionalities such as banner management, publicity management, visitor monitoring, user management, news uploads, and gallery updates. These features enable administrators to control content and monitor platform activity effectively. Additionally, the analytics module generates graphical representations such as property demand trends, pricing analysis, regional performance, and user engagement statistics. The data visualizations assist stakeholders in making informed strategic decisions regarding marketing strategies, pricing policies, and investment planning. By transforming aggregated data into actionable intelligence, the BI module enhances transparency, operational efficiency, and business growth within the smart real estate ecosystem.

Performance Analysis and Discussion

The proposed Privacy-Aware Federated Learning Framework integrated with the Smart Living Real Estate Portal and Business Intelligence (BI) module was evaluated across multiple performance dimensions, including recommendation accuracy, privacy protection strength, communication efficiency, scalability, and analytics performance. The system was tested in a simulated distributed environment representing multiple smart habitation clients participating in collaborative model training.From an accuracy perspective, the federated learning model achieved a recommendation accuracy of 91.7%, outperformingthe traditional centralized model, which recorded 88.4%. This improvement is attributed to localized training on user-specific behavioural patterns while still benefiting from aggregated global knowledge. The distributed architecture enables personalized learning without compromising model generalization.

In terms of privacy protection, the framework significantly reduced data leakage risks by ensuring that raw user data never leaves local devices. The integration of differential privacy added controlled statistical noise to model updates, preventing inference attacks. Secure aggregation protocols and encryption techniques further strengthened confidentiality during communication and aggregation. Compared to centralized systems where user data is stored in a single repository, the proposed model demonstrated a substantially lower exposure to privacy threats.

Communication overhead was optimized through selective parameter sharing and efficient aggregation rounds. Although cryptographic mechanisms such as Secure Multi-Party Computation (SMPC) and Homomorphic Encryption introduced moderate computational overhead, the distributed workload balanced processing demands across client devices, reducing server-side burden and enhancing scalability. The Business Intelligence module successfully transformed aggregated outputs into actionable insights. The system achieved 89.5% accuracy in regional demand forecasting and 90.8% accuracy in price trend prediction. Dashboard response time remained within 1.8 seconds, ensuring near real-time visualization and analytics delivery. The BI analytics supported stakeholders in strategic decision-making, targeted advertising, and market positioning.

Overall, the experimental evaluation confirms that the proposed framework provides a secure, scalable, and efficient solution for smart habitation ecosystems. The model achieves a balance between analytical performance and strong privacy preservation, making it suitable for real-world deployment in privacy-sensitive real estate and smart living environments.

TABLE I

Category	Metric	Centralized Approach	Proposed Federated Framework	Observation
Model Performance	Recommendation Accuracy	88.4%	91.7%	Improved personalization
Privacy Protection	Raw Data Sharing	Yes	No	Strong privacy guarantee
Privacy Risk	Data Leakage Probability	High	Very Low	Risk minimized
Communication	Overhead	Moderate	Low-Moderate	Optimized updates
Computation	Server Load	High	Distributed	Balanced workload
Scalability	Multi-Client Support	Limited	High	Easily scalable
BI Analytics	Demand Forecast Accuracy	85%	89.5%	Better market insights
BI Analytics	Price Trend Prediction	87%	90.8%	Improved forecasting
System Efficiency	Dashboard Response Time	2.5 sec	1.8 sec	Faster analytics

CONCLUSION

This study proposed a privacy-aware federated learning framework for secure data collaboration within a smart living real estate ecosystem. The system was designed to eliminate the need for centralized raw data storage by enabling local model training on user devices while sharing only protected model updates for global aggregation. By integrating privacy-enhancing techniques such as differential privacy, secure aggregation, Secure Multi-Party Computation, and homomorphic encryption, the framework significantly reduces the risk of sensitive information leakage. Experimental evaluation demonstrated improved recommendation accuracy, balanced computational load, and enhanced scalability compared to conventional centralized approaches. Furthermore, the integration of a Business Intelligence module enables meaningful analytics, including demand forecasting and price trend analysis, thereby supporting informed decision-making for stakeholders. Overall, the proposed architecture successfully achieves a balance between intelligent data utilization and strong privacy preservation, making it suitable for deployment in modern smart habitation environments.

REFERENCES

[1.] Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., & Zhang, L. (2016). Deep learning with differential privacy. *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, 308–318.

[2.] Aledhari, M., Razzak, R., Parizi, R. M., & Srivastava, G. (2020). Federated learning: A survey on enabling technologies, protocols, and applications. *IEEE Access, 8*, 140699–140725.

[3.] Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2017). Practical secure aggregation for privacy-preserving machine learning. *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, 1175–1191.

[4.] Chen, H., Chiang, R. H. L., &Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165–1188.

[5.] Dwork, C. (2006). Differential privacy. *Proceedings of the International Colloquium on Automata, Languages and Programming*, 1–12.

[6.] Dwork, C., & Roth, A. (2014). *The algorithmic foundations of differential privacy*.

[7.] Evans, D., Kolesnikov, V., &Rosulek, M. (2018). A pragmatic introduction to secure multi-party computation. *Foundations and Trends in Privacy and Security*, *2*(2–3), 70–246.

[8.] Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144.

[9.] Gentry, C. (2009). Fully homomorphic encryption using ideal lattices. *Proceedings of the ACM Symposium on Theory of Computing*, 169–178.

[10.] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[11.] Kimball, R., & Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling* (3rd ed.). Wiley.

[12.] Konečný, J., McMahan, H. B., Yu, F. X., Richtárik, P., Suresh, A. T., & Bacon, D. (2016). Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*.

[13.] Li, S., Xu, L. D., & Zhao, S. (2019). Secure and privacy-preserving data aggregation in IoT. *IEEE Internet of Things Journal, 6*(3), 5657–5666.
[14.] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems*, 429–450.
[15.] Lindell, Y. (2021). Secure multiparty computation. *Communications of the ACM, 64*(1), 86–96.
[16.] Liu, X., Cao, J., Tang, S., & Wen, J. (2019). Edge computing for IoT-based smart homes. *IEEE Network, 33*(2), 90–96.
[17.] Lyu, L., Yu, H., Yang, Q., & Jin, Y. (2020). Threats to federated learning: A survey. *arXiv preprint arXiv:2003.02133*.
[18.] McMahan, B., Moore, E., Ramage, D., Hampson, S., &Arcas, B. A. y. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of the International Conference on Artificial Intelligence and Statistics*, 1273–1282.
[19.] McMahan, H. B., Ramage, D., Talwar, K., & Zhang, L. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning, 14*(1–2), 1–210.
[20.] Modak, N. M., Panda, S., & Sana, S. S. (2019). IoT-enabled smart cities: A review. *Sustainable Cities and Society, 50*, 101678.
[21.] Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system*. https://bitcoin.org/bitcoin.pdf
[22.] O'Reilly, T. (2007). What is Web 2.0: Design patterns and business models for the next generation of software. *Communications & Strategies, 65*, 17–37.
[23.] Paillier, P. (1999). Public-key cryptosystems based on composite degree residuosity classes. *Proceedings of EUROCRYPT*, 223–238.
[24.] Rahman, A., Rahman, S., & Rashid, M. M. (2018). Privacy-preserving smart home data analytics. *IEEE Internet of Things Journal, 5*(6), 4756–4767.
[25.] Shokri, R., & Shmatikov, V. (2015). Privacy-preserving deep learning. *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, 1310–1321.
[26.] Swan, M. (2015). *Blockchain: Blueprint for a new economy*. O'Reilly Media.
[27.] Voigt, P., & von dem Bussche, A. (2017). *The EU General Data Protection Regulation (GDPR): A practical guide*. Springer.
[28.] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology, 10*(2), 1–19.
[29.] Yu, W., Liu, S., Wang, X., & Zhou, X. (2022). A survey on federated learning systems. *IEEE Communications Surveys & Tutorials, 24*(1), 1–36