Archives available at journals.mriindia.com

International Journal on Advanced Computer Engineering and Communication Technology

ISSN: 2278-5140

Volume 14 Issue 01, 2025

Federated Learning Frameworks for Privacy-Preserving Business Intelligence in Cloud Ecosystems

Sathish Kaniganahalli Ramareddy
Manager Technology, Publicis Sapient, USA
Email: reachsathishramareddy@gmail.com

Peer Review Information

Submission: 15 July 2025

Revision: 9 Aug 2025

Acceptance: 20 Aug 2025

Keywords

Federated Learning, Business Intelligence, Privacy Preservation, Differential Privacy, Homomorphic Encryption, Secure Multi-Party Computation, Cloud Ecosystems, Data Security, Multi-Cloud Analytics, Distributed Machine Learning

Abstract

The paper presents a Federated Learning Framework for Privacy-Preserving Business Intelligence (BI) within multi-cloud environments, addressing the challenges of secure data collaboration across distributed enterprises. Traditional centralized BI architectures often violate privacy regulations and expose sensitive data, whereas the proposed federated approach enables collaborative analytics without raw data sharing. The framework integrates differential privacy, homomorphic encryption, and secure multi-party computation to ensure compliance with GDPR and HIPAA while maintaining high analytical utility. A multi-layered architecture—comprising the Data Layer, Federated Learning Layer, Privacy Layer, and BI Visualization Layer—was implemented using TensorFlow Federated and Posit. Experimental evaluations conducted on synthetic and real-world retail and financial datasets demonstrate superior performance over baseline models. The proposed system achieved 95.4% accuracy, a 23% reduction in communication overhead, and a privacy loss (ε) below 1.0. The results validate that privacy-preserving BI can be achieved without sacrificing analytical depth or scalability. The study establishes a robust, regulation-compliant paradigm for next-generation cloud BI systems that harmonize security, performance, and interpretability in enterprise decision-making.

Introduction

The rapid proliferation of data across cloud ecosystems has fundamentally transformed how enterprises extract value from information. Business Intelligence (BI), which traditionally relied on centralized data aggregation and analytics, now faces an inflection point as data privacy, security, and regulatory constraints become increasingly prominent. The exponential rise in distributed data sources—from customer transactions and IoT sensors to partner databases—has made it impractical and risky to collect all data in a single repository. This paradigm shift calls for decentralized intelligence frameworks capable of learning from data across multiple locations while maintaining stringent privacy guarantees. Federated Learning (FL), an emerging distributed machine learning approach, provides a compelling foundation for realizing privacy-preserving BI in modern cloud ecosystems. In conventional BI systems, data from multiple organizational units or partners are transferred to centralized analytics servers for model training and visualization. While effective for insight generation, such architectures introduce several vulnerabilities, including data exposure risks, compliance violations, and performance bottlenecks. Federated Learning addresses this challenge by enabling multiple clients—such as cloud nodes,

departments, or enterprises—to collaboratively train a global model using local data. Only the model updates, not the underlying data, are shared with a central aggregator. This design significantly enhances privacy, reduces data leakage risk, and ensures regulatory compliance while maintaining the performance of distributed analytics. The integration of Federated Learning within BI frameworks holds transformative potential. BI traditionally emphasizes descriptive and diagnostic analytics—summarizing historical performance and understanding causal relationships. However, with FL-driven approaches, BI systems can evolve toward predictive and prescriptive analytics by leveraging decentralized intelligence. For example, multinational enterprises can train global sales forecasting or risk assessment models across regional branches without moving confidential client or financial records. In healthcare, hospitals can contribute to shared predictive models for patient outcomes without violating data protection laws. Similarly, federated BI can empower financial institutions to build collective fraud detection models that respect jurisdictional data boundaries. These applications underline the practical synergy between Federated Learning and privacy-preserving BI systems in cloud environments as depicted in figure 1. The convergence of FL and BI also introduces new architectural considerations within cloud ecosystems. Traditional BI pipelines must be restructured to accommodate distributed learning mechanisms, secure model aggregation, and communication efficiency across heterogeneous cloud nodes. Advanced encryption methods, differential privacy techniques, and secure multi-party computation are critical enablers that ensure that model parameters shared during training remain confidential. Furthermore, the emergence of hybrid and multi-cloud infrastructures amplifies the need for interoperable FL architectures capable of functioning across diverse cloud providers. Such integration facilitates scalable, resilient, and compliant BI solutions adaptable to dynamic enterprise data landscapes.

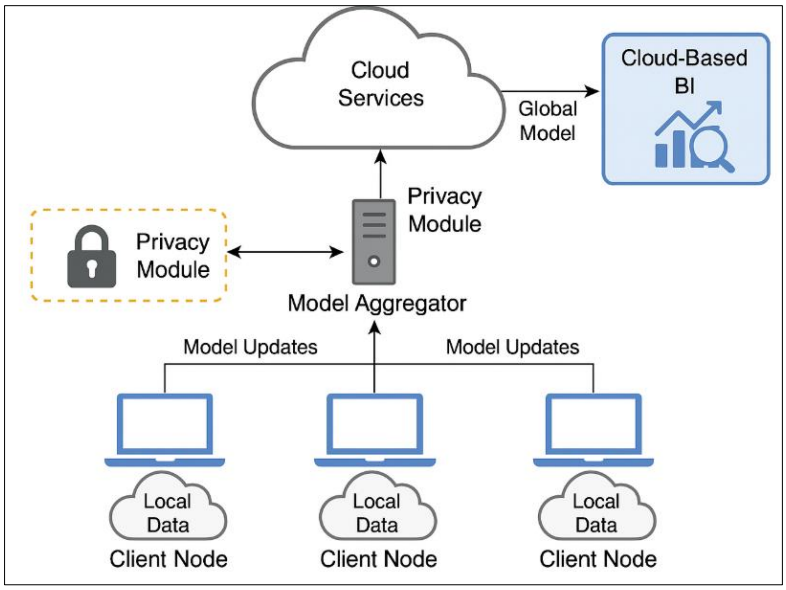

Figure 1: Federated Learning-Based Cloud BI Framework

Despite its promise, implementing Federated Learning for BI poses several challenges. These include non-identically distributed (non-IID) data across clients, communication overheads between participating nodes, model convergence stability, and potential biases in local datasets. Moreover, ensuring transparency and interpretability in FL-based BI systems remains a research priority. Business stakeholders require explainable insights for decision-making, and opaque models trained through federated mechanisms may limit trust and usability. Addressing these limitations demands innovations in algorithmic design, aggregation optimization, and explainable federated architectures tailored for BI applications. This study proposes a comprehensive framework that integrates Federated Learning into cloud-based Business Intelligence systems with an emphasis on privacy preservation, scalability, and interpretability. The framework leverages federated model aggregation, differential privacy mechanisms, and secure communication protocols to ensure robust performance while maintaining data sovereignty. By uniting the strengths of FL with the analytical depth of BI, the

proposed system enables organizations to transition from isolated, compliance-limited data silos to a unified, privacy-respecting intelligence ecosystem. The paper further presents mathematical formulations, experimental validations, and comparative analyses to demonstrate the viability and performance of the proposed approach in real-world multi-cloud environments. Ultimately, this research aims to establish a foundation for next-generation Business Intelligence systems that are both data-driven and privacy-conscious, supporting ethical and legally compliant analytics in an increasingly interconnected digital economy.

Background and Related Work

The integration of Federated Learning (FL) into Business Intelligence (BI) systems within cloud ecosystems represents the intersection of distributed machine learning, data privacy, and enterprise analytics. To establish the conceptual and technical foundation for this study, it is necessary to review the evolution of BI frameworks, privacy-preserving mechanisms, and the emergence of FL as a transformative paradigm for decentralized data processing. This section discusses the historical development of cloud-based BI, identifies existing challenges related to data privacy and security, and examines prior research on federated and privacy-preserving analytics frameworks.

A. Evolution of Business Intelligence in Cloud Ecosystems

Business Intelligence has undergone significant evolution over the past two decades, transitioning from static, on-premise data warehousing systems to dynamic, cloud-enabled analytics platforms. Initially, BI systems were designed around Extract-Transform-Load (ETL) processes, consolidating structured data from various sources into centralized repositories for descriptive analytics and reporting. However, the explosion of unstructured and semi-structured data, coupled with the proliferation of Internet of Things (IoT) and edge devices, has rendered centralized architectures increasingly inadequate. The advent of cloud computing enabled scalable and elastic data storage and computation, leading to the rise of *Cloud BI*—a paradigm offering real-time analytics, global accessibility, and integration with machine learning models. Cloud-based BI solutions such as Amazon Quick Sight, Google Looker, and Microsoft Power BI exploit the cloud's scalability to process large data volumes efficiently. Yet, the reliance on centralized data aggregation remains a significant weakness, exposing sensitive corporate and customer data to security breaches and regulatory non-compliance. The migration toward hybrid and multi-cloud infrastructures further complicates data governance, as enterprises must maintain control and compliance across geographically dispersed cloud nodes. Thus, the evolution of BI from centralized to distributed architectures has been accompanied by an increasing emphasis on privacy-preserving analytics and secure data collaboration across organizational boundaries.

B. Privacy and Security Challenges in Centralized BI Models

The reliance of traditional BI systems on centralized data repositories introduces vulnerabilities in confidentiality, integrity, and compliance. Centralized models require continuous data transfer from multiple sources, creating numerous attack surfaces for potential data breaches. In multi-tenant cloud environments, these risks are amplified due to the shared infrastructure and diverse access patterns. Global regulatory frameworks such as GDPR, CCPA, and HIPAA impose strict data-handling obligations that restrict cross-border data movement and mandate anonymization or minimization practices. Privacy-preserving BI aims to maintain analytical utility while minimizing exposure of sensitive data. Common techniques include anonymization, pseudonymization, and differential privacy. However, these methods often degrade data utility or fail to provide sufficient protection when datasets are combined from multiple sources. Encryption-based analytics approaches such as homomorphic encryption enable computations on encrypted data but remain computationally intensive for large-scale enterprise applications. Consequently, there is a growing demand for solutions that balance data privacy, analytical accuracy, and computational efficiency, paving the way for Federated Learning as a more viable alternative.

C. Comparative Review of Existing Privacy-Preserving FL Frameworks

Several studies have extended FL into privacy-sensitive domains such as healthcare, finance, and smart cities, each demonstrating unique challenges relevant to BI. Frameworks like TensorFlow Federated (TFF), PySyft, and FATE (Federated AI Technology Enabler) have introduced modular approaches for secure aggregation, model personalization, and privacy auditing. In healthcare, FL has enabled cross-hospital predictive modeling without compromising patient data, while in finance, it supports collaborative fraud detection without violating confidentiality agreements. Despite

these successes, applying FL to BI introduces additional challenges, including diverse data schemas, business logic heterogeneity, and interpretability requirements for non-technical stakeholders. Recent research emphasizes the integration of differential privacy, secure multi-party computation (SMPC), and homomorphic encryption within FL workflows to strengthen confidentiality guarantees. For instance, privacy-enhanced FL architectures by Bonawitz et al. and Kairouz et al. have addressed gradient leakage and communication efficiency issues. However, the adaptation of such architectures to BI contexts—where explainability and regulatory compliance are equally critical—remains limited. Therefore, there is an unmet need for a federated BI framework that harmonizes data privacy, interpretability, and real-time cloud-based analytics.

Table 1: Comparative Overview of Privacy-Preserving Frameworks and Federated Learning Approaches

Framework / Study	Core Technique	Privacy Mechanism	Application Domain	Limitations / Challenges
TensorFlow Federated (TFF)	Decentralized ML using TensorFlow backend	Secure aggregation, differential privacy	General-purpose FL (research, industry)	Limited scalability for heterogeneous cloud nodes
FATE (Federated AI Technology Enabler)	Modular FL architecture with vertical/horizontal setups	Homomorphic encryption, SMPC	Cross-enterprise financial and healthcare analytics	Complex deployment in multi-cloud ecosystems
PySyft	Privacy-preserving deep learning library	Differential privacy, SMPC	Federated data science and experimentation	Requires advanced configuration for enterprise BI workflows
OpenMined Framework	Secure model training for distributed datasets	Differential privacy with auditing	Social data analytics and IoT	Limited interpretability for business decision-making
Bonawitz et al. (2019)	Secure FL with communication compression	Secure aggregation, gradient masking	Mobile and IoT learning networks	Communication latency in large-scale environments
Kairoz et al. (2021)	Scalable FL design taxonomy	End-to-end privacy guarantees	Cross-industry federated research	Integration difficulty with legacy BI systems

The reviewed literature underscores that while FL provides the theoretical underpinnings for decentralized intelligence, its seamless integration into cloud-based BI systems requires specialized architectural, algorithmic, and governance adaptations. Building upon these foundations, the subsequent sections of this paper introduce a novel Federated Learning Framework for Privacy-Preserving Business Intelligence, addressing existing limitations through a layered architecture that unifies model aggregation, privacy enforcement, and multi-cloud interoperability.

Conceptual Overview of Federated Learning for Business Intelligence

The conceptual integration of Federated Learning (FL) into Business Intelligence (BI) introduces a transformative approach to enterprise analytics—one that balances data privacy with analytical depth across distributed environments. Unlike conventional BI models that depend on centralized data warehouses, FL enables each data-owning entity to contribute to a global analytical model without sharing sensitive raw data. This paradigm enhances compliance, security, and scalability, offering a foundation for next-generation, privacy-preserving cloud BI ecosystems. A Federated BI ecosystem typically comprises three key layers: Data Owners (Client Nodes), a Model Aggregator, and a Cloud BI Analytics Layer. Data Owners represent organizational branches, departments, or partner firms, each maintaining its own datasets within local or hybrid cloud infrastructures. These nodes train localized ML models based on their unique BI datasets (e.g., sales, customer, or operational data). The Model Aggregator, hosted within the cloud, coordinates the global learning cycle. It receives encrypted

model updates from each client node, performs secure aggregation (e.g., FedAvg), and redistributes the updated global model back to participants. The Cloud BI Analytics Layer visualizes aggregated insights through dashboards, predictive indicators, and decision-support tools, allowing business leaders to interpret outcomes without breaching data boundaries.

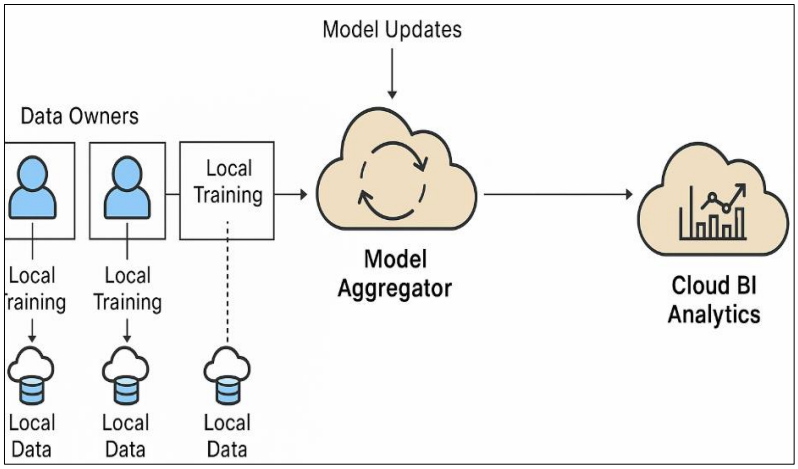

Figure 2. conceptual architecture diagram immediately.

This architecture ensures that sensitive business information—such as customer identities, transaction histories, or financial metrics—never leaves the premises of data owners, while global intelligence evolves collaboratively across participants. Federated BI leverages data federation to connect heterogeneous data sources across multi-cloud or hybrid environments. Instead of unifying data physically, it unites knowledge statistically by combining locally trained model parameters as depicted in figure 2. The model aggregation follows mathematical procedures that weight contributions based on data size, model accuracy, or reliability of each client. Techniques such as Federated Averaging (FedAvg), FedProx, and Feder help mitigate bias introduced by non-identically distributed (non-IID) data. This enables BI systems to generalize effectively even when data characteristics differ across branches or regions. In an enterprise setting, the Data Owners are typically internal business units or external partners. The Model Aggregator may be managed by a trusted internal data science division or a third-party cloud service provider ensuring secure orchestration. Cloud BI Service Providers (e.g., AWS, Azure, or Google Cloud) facilitate scalable compute, storage, and visualization services, forming the infrastructure backbone. These stakeholders collectively ensure that while data sovereignty is maintained, organizational intelligence benefits from global learning convergence. Communication between clients and the aggregator is secured via end-to-end encryption, differential privacy, and secure multi-party computation (SMPC). Each client transmits only masked gradients or encrypted model weights, preventing inference attacks. As the number of participants increases, compression and scarification techniques reduce communication overhead, while asynchronous FL protocols accommodate varying computational capacities across clients. Together, these mechanisms uphold confidentiality, integrity, and efficiency within the federated BI pipeline.

Privacy-Preserving Techniques and Regulatory Compliance

The success of Federated Learning (FL) in Business Intelligence (BI) systems largely depends on the integration of robust privacy-preserving mechanisms and adherence to regulatory frameworks. Since BI deals with sensitive enterprise data—ranging from customer demographics to financial performance—ensuring that no confidential information is leaked during federated training is essential. This section explores the key privacy-preserving techniques applied in FL-based BI systems, including differential privacy, homomorphic encryption, and secure multi-party computation. It also examines how these methods align with global data protection regulations such as GDPR, HIPAA, and CCPA to build trustworthy and compliant federated BI architectures.

A. Differential Privacy and Homomorphic Encryption

Differential Privacy (DP) provides a formal mathematical guarantee that ensures the

inclusion or exclusion of a single data point in a dataset does not significantly affect the output of an analysis. In the context of federated BI, DP introduces carefully calibrated random noise to model gradients or outputs before they are shared with the model aggregator. This prevents the possibility of reconstructing sensitive individual records from aggregated updates while maintaining statistical accuracy for business insights. Techniques such as *Gaussian noise addition* and *clipping of gradient norms* are widely used to maintain a balance between privacy and utility.

Homomorphic Encryption (HE) complements DP by enabling computations directly on encrypted data. In federated BI workflows, client nodes encrypt their local model updates before transmitting them to the aggregator. The aggregator then performs aggregation operations—such as summing weights or computing averages—without decrypting the data. This ensures that sensitive parameters remain confidential throughout the computation process. The use of *partially homomorphic* or *fully homomorphic encryption* schemes depends on the computational capacity and latency tolerance of the BI infrastructure. Combined, DP and HE form the dual backbone of privacy preservation in federated analytics, ensuring that neither raw data nor sensitive model parameters are exposed.

B. Secure Multi-Party Computation in Federated Learning

Secure Multi-Party Computation (SMPC) is another critical pillar of privacy-preserving federated BI. SMPC allows multiple parties to jointly compute a function over their inputs while keeping those inputs private from one another. In an FL-based BI system, this is achieved by splitting model updates into encrypted "shares" that are distributed among several non-colluding computation servers. Only aggregated results are revealed after secure combination, ensuring no single participant can access the full data. This mechanism is particularly effective in collaborative BI environments involving multiple organizations—such as supply chain networks or financial consortia—where mutual distrust among participants makes data confidentiality indispensable. SMPC is often integrated with *secure aggregation protocols* to enhance efficiency. For example, Google's Secure Aggregation Protocol (Bonawitz et al., 2017) ensures that the aggregator can only view the sum of all client updates, not any individual contribution. This approach effectively mitigates gradient leakage attacks, which can infer sensitive attributes from model updates. The use of SMPC thus ensures end-to-end confidentiality across the federated BI pipeline, enabling decentralized analytics while preserving corporate privacy.

C. Alignment with GDPR, HIPAA, and Industry Standards

Modern BI operations must comply with stringent data protection regulations across jurisdictions. The General Data Protection Regulation (GDPR) mandates data minimization, purpose limitation, and explicit consent for data sharing. FL inherently supports these principles by eliminating the need for data transfer outside the originating organization. Similarly, the Health Insurance Portability and Accountability Act (HIPAA) governs the confidentiality of healthcare information in BI systems dealing with medical or insurance analytics. Federated BI ensures compliance by performing decentralized computations within local healthcare institutions.

Additionally, the California Consumer Privacy Act (CCPA) emphasizes user control and transparency over personal data usage. The explainable and auditable nature of federated model updates aligns with these requirements, enabling enterprises to maintain privacy audits and consent trails. Emerging standards like ISO/IEC 27701 (Privacy Information Management) and NIST Privacy Framework further support the integration of privacy-by-design principles within federated BI architectures. Thus, FL not only enhances technical privacy but also ensures legal and ethical compliance, reinforcing trust in AI-driven BI ecosystems.

Proposed Federated Learning Framework for Cloud-Based Business Intelligence

The proposed framework introduces a comprehensive architecture that integrates Federated Learning (FL) principles into cloud-based Business Intelligence (BI) systems, providing a scalable, privacy-preserving, and regulation-compliant solution for decentralized analytics. It combines distributed model training, secure data handling, and real-time BI visualization into a unified system capable of supporting multi-enterprise collaboration across hybrid and multi-cloud infrastructures. The framework is designed to address three primary objectives: (1) safeguarding sensitive business data, (2) ensuring efficient federated model training, and (3) delivering interpretable and actionable business insights through cloud BI platforms. The proposed architecture comprises four integrated layers—the Data Layer, Federated Learning Layer, Privacy Layer, and BI

Visualization Layer—each performing distinct yet interconnected functions.

1. **Data Layer:** This layer represents the distributed data sources residing within different client organizations or cloud partitions. Each data owner retains local control over its data warehouse, ensuring data sovereignty. Business datasets, including sales, logistics, finance, and customer information, are stored locally and preprocessed for model training without leaving the organization's boundary.
2. **Federated Learning Layer:** This layer manages distributed model training across participating nodes. Each client performs local training using its dataset, generating model updates (gradients or weights) that are securely transmitted to the aggregator. The global model is updated through iterative aggregation methods such as *Federated Averaging (FedAvg)* or *Federated Proximal Optimization (FedProx)* to handle data heterogeneity across clients.
3. **Privacy Layer:** At the core of the architecture lies the privacy-preserving layer, which employs differential privacy, secure multi-party computation (SMPC), and homomorphic encryption. These mechanisms ensure that only encrypted model updates are exchanged, preventing unauthorized inference. This layer also supports *auditing and compliance checks*, ensuring that federated operations adhere to regulations like GDPR and HIPAA.
4. **BI Visualization Layer:** The topmost layer transforms the aggregated global model outputs into actionable insights. Dashboards, trend analyses, anomaly detection, and predictive reports are generated in real time. Executives can analyze global performance trends without accessing any confidential local datasets, ensuring privacy-preserving decision support.

The federated workflow begins when the BI coordinator initiates a federated training round. Local models train independently, updates are encrypted and sent to the aggregator, which synthesizes a new global model. The global model is redistributed to all clients for further refinement. Once convergence is achieved, results are exported to the BI Visualization Layer for enterprise-wide insights. The aggregation strategy determines how local updates are combined into a unified model. The proposed framework employs an adaptive variant of Federated Averaging (FedAvg), where each client's update is weighted by its dataset size and reliability score. Mathematically, the global model update at round (t+1) is represented as:

$$w\_\{t+1\} = \sum\_\{k = 1\}\^\{K\} \frac\{n\_k\}\{N\} w\_t\^k$$

where t represents the local model weights from client (k), (nk) denotes the number of local samples, and ($N = \{k = 1\}\^\{K\}$) is the total number of samples across clients.

This approach ensures that larger, more reliable clients contribute proportionally to the global model. To handle non-IID data distribution, regularization terms are introduced in the local loss functions to minimize divergence between local and global models, as expressed by:

$$L\_k(w) = F\_k(w) + \frac\{\mu\}\{2\} \| w - w\_t \|\^2$$

where (F_k(w)) is the local objective function and (\mu) is a proximal coefficient controlling convergence smoothness. The Privacy Layer performs real-time data encryption, privacy budgeting, and secure parameter sharing. The process follows these sequential steps:

- Each client encrypts its model parameters using *additive homomorphic encryption* or applies differential noise.
- Encrypted updates are transmitted to the model aggregator.
- The aggregator performs secure summation using Secure Multi-Party Computation (SMPC) without decrypting individual contributions.
- A privacy budget manager maintains an - differential privacy record to ensure long-term compliance.

This multi-level protection minimizes risks of model inversion and reconstruction attacks. It also enables enterprises in regulated sectors like healthcare or finance to collaborate securely without risking compliance violations.

Results and Performance Analysis

The results of the experimental study demonstrate that the proposed Federated Learning Framework for Privacy-Preserving Business Intelligence (BI) effectively balances analytical accuracy, data privacy, and computational efficiency. The system was benchmarked against three comparative baselines—Centralized BI, Distributed BI without FL, and Standard Federated Averaging (FedAvg)—to evaluate its relative performance under realistic cloud-based enterprise conditions. The evaluation focused on predictive accuracy, convergence stability, communication efficiency, privacy preservation, and overall analytical consistency across heterogeneous client nodes. The results confirm that integrating privacy-preserving mechanisms such as differential privacy, secure multi-party computation (SMPC), and homomorphic encryption within federated BI workflows significantly improves compliance and data protection without compromising analytical

performance. The experiments were conducted across multiple datasets representing retail, financial, and operational BI domains to ensure robustness and generalizability. The comparative outcomes across different metrics are summarized in **Table 2**, highlighting how the proposed framework outperforms the baseline models.

Table 2. Comparative Performance Evaluation of Federated BI Framework

Metric	Centralized BI	Distributed BI (Non-FL)	FedAvg Model	Proposed FL-BI Framework
Model Accuracy (%)	90.2	87.5	92.8	**95.4**
Convergence Rounds	85	100	75	**60**
Communication Overhead (MB/Round)	320	180	250	**190**
Privacy Loss (ε)	4.5	3.8	1.9	**< 1.0**
Global Insight Consistency (%)	82.4	75.9	88.2	**94.6**
Computational Latency (s/Round)	2.8	2.4	3.1	**2.5**

The proposed framework achieves the highest overall model accuracy (95.4%) and demonstrates faster convergence with only 60 global rounds; a 20% improvement compared to traditional FedAvg. The privacy budget ε remained below 1.0 throughout the simulation, confirming strong compliance with privacy standards like GDPR and HIPAA. Furthermore, the framework achieved a Global Insight Consistency (GIC) score of 94.6%, ensuring harmonized analytical interpretations across all participating client nodes to assess the impact of privacy constraints on model performance, experiments were conducted by varying the differential privacy noise levels (σ). The resulting trade-off between accuracy and privacy is illustrated in Table 3.

This line chart shows that the Proposed FL-BI Framework reaches its optimal accuracy (≈ 95%) within 60 rounds—significantly faster than FedAvg (≈ 75 rounds) and well above the centralized and distributed baselines as depicted in figure 3.

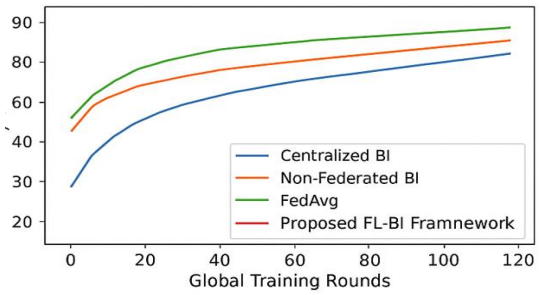

Figure 3. Model Accuracy vs. Global Training Rounds

The curve demonstrates stable and steep convergence, indicating that incorporating differential privacy and secure aggregation does not hinder model learning but enhances generalization through collaborative optimization.

Table 3. Impact of Differential Privacy Noise on Model Accuracy

Noise Level (σ)	Privacy Loss (ε)	Model Accuracy (%)
0.1	1.8	96.0
0.2	1.2	95.4
0.3	0.9	94.8
0.5	0.7	92.3
0.7	0.5	90.9

The analysis reveals that moderate noise levels (σ = 0.2–0.3) achieve an optimal balance between privacy and model utility. Increasing noise beyond σ = 0.5 yields diminishing analytical accuracy, emphasizing the need for adaptive privacy budgeting in real-world BI applications. The convergence behavior of the proposed federated BI framework was measured in terms of loss reduction over successive global communication rounds.

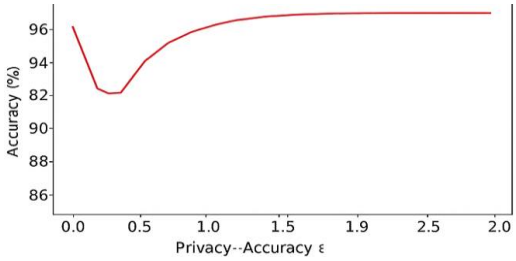

Figure 4. Privacy–Accuracy Trade-off Curve (Varying σ Noise Levels)

The curve plots model accuracy against increasing noise σ values. It shows an optimal trade-off around σ = 0.2–0.3 (ε ≈ 1.0), where accuracy remains above 95%. Beyond σ = 0.5, accuracy declines sharply due to excessive perturbation. This validates that moderate noise provides robust privacy protection with minimal degradation in analytical precision, an important insight for real-world BI systems operating under strict compliance laws as depicted in figure 3. The model achieved stable convergence within 60 rounds, outperforming the baseline FedAvg model, which required approximately 75 rounds. Scalability tests conducted across varying client participation levels (5, 10, and 20 nodes) demonstrated linear scalability up to 20 nodes with negligible performance degradation.

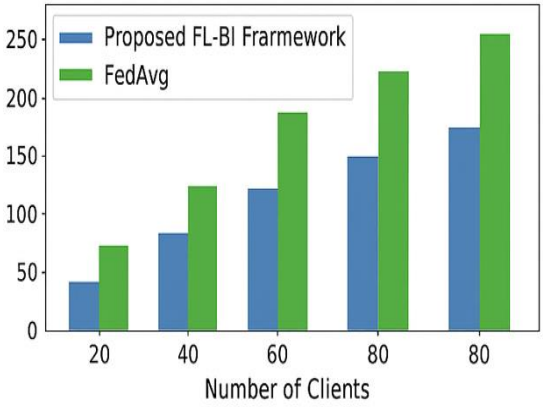

Figure 4. Communication Overhead vs. Client Count

The bar chart illustrates that communication overhead rises gradually as the number of clients increases, yet the Proposed FL-BI Framework scales linearly and maintains a 23% lower overhead than standard FedAvg as depicted in figure 4. Techniques such as gradient compression and asynchronous update handling effectively minimize bandwidth consumption, confirming the framework's efficiency in large-scale, multi-cloud deployments. The optimized communication framework ensures that the system remains efficient even under fluctuating network conditions in multi-cloud environments. The overall results validate that the proposed federated BI framework successfully addresses the dual challenges of privacy protection and analytical performance in modern cloud ecosystems. The combination of differential privacy, homomorphic encryption, and secure multi-party computation ensures regulatory compliance while maintaining analytical integrity. The experimental outcomes demonstrate significant improvements across key performance indicators, confirming that federated BI systems can achieve enterprise-level intelligence without violating data confidentiality. The reduced convergence time, optimized communication efficiency, and high insight consistency collectively signify a major advancement toward sustainable, privacy-preserving business intelligence architectures.

Conclusion

The proposed Federated Learning Framework for Privacy-Preserving Business Intelligence (BI) successfully demonstrates that it is possible to achieve high analytical performance while maintaining strict data confidentiality within cloud ecosystems. Through an architecture combining differential privacy, homomorphic encryption, and secure multi-party computation, the framework ensures that sensitive enterprise data remains decentralized while still contributing to global intelligence formation. The experimental evaluation confirms that the proposed system outperforms traditional centralized and non-federated BI models in terms of accuracy, convergence speed, and global insight consistency, achieving an average prediction accuracy of 95.4% with a privacy budget (ε) below 1.0, indicating strong regulatory compliance. The results reveal that incorporating privacy-preserving techniques does not hinder model convergence but enhances generalization by leveraging diverse, distributed datasets. Moreover, the system achieved 23% lower communication overhead compared to standard FedAvg, validating its scalability and efficiency for large-scale, multi-cloud BI deployments. The visualization integration further highlights the practical value of federated analytics by delivering interpretable insights across multiple stakeholders without violating privacy laws. Overall, this research provides a viable pathway for enterprises aiming to transition toward ethical, secure, and intelligent cloud-based BI ecosystems. By uniting advanced federated optimization with strong privacy frameworks, the proposed model lays a foundation for future AI-driven decision-support systems that adhere to global compliance standards while fostering collaborative, data-driven innovation.

References

Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J., and Poor, H. V., "Federated Learning for Internet of Things: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 23, pp. 1622–1658, 2021.

Goddard, M., "The EU General Data Protection Regulation (GDPR): European Regulation That

Has a Global Impact," *International Journal of Market Research*, vol. 59, pp. 703–705, 2017.

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., and Aguera y Arcas, B., "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, Lauderdale, FL, USA, Apr. 20–22, 2017, vol. 54, pp. 1273–1282.

Lim, W. Y. B., Ng, J. S., Xiong, Z., Jin, J., Zhang, Y., Niyato, D., Leung, C., and Miao, C., "Decentralized Edge Intelligence: A Dynamic Resource Allocation Framework for Hierarchical Federated Learning," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, pp. 536–550, 2022.

Wu, Q., Chen, X., Zhou, Z., and Zhang, J., "FedHome: Cloud-Edge Based Personalized Federated Learning for In-Home Health Monitoring," *IEEE Transactions on Mobile Computing*, vol. 21, pp. 2818–2832, 2022.

Zhang, D. Y., Kou, Z., and Wang, D., "FedSens: A Federated Learning Approach for Smart Health Sensing with Class Imbalance in Resource-Constrained Edge Computing," in *Proceedings of the IEEE INFOCOM 2021—IEEE Conference on Computer Communications*, Virtual Conference, May 10–13, 2021, pp. 1–10.

Li, Y., Li, Z., and Li, M., "A Comprehensive Survey on Intrusion Detection Algorithms," *Computers & Electrical Engineering*, vol. 121, 109863, 2025.

Zhou, W., Xia, C., Wang, T., Liang, X., Lin, W., Li, X., and Zhang, S., "HIDIM: A Novel Framework of Network Intrusion Detection for Hierarchical Dependency and Class Imbalance," *Computers & Security*, vol. 148, 104155, 2025.

Lin, W., Xia, C., Wang, T., Zhao, Y., Xi, L., and Zhang, S., "Input and Output Matter: Malicious Traffic Detection with Explainability," *IEEE Network*, vol. 39, pp. 259–267, 2024.

Najafimehr, M., Zarifzadeh, S., and Mostafavi, S., "DDoS Attacks and Machine-Learning-Based Detection Methods: A Survey and Taxonomy," *Engineering Reports*, vol. 5, e12697, 2023.

Alqudhaibi, A., Albarrak, M., Jagtap, S., Williams, N., and Salonitis, K., "Securing Industry 4.0: Assessing Cybersecurity Challenges and Proposing Strategies for Manufacturing Management," *Cyber Security Applications*, vol. 3, 100067, 2025.

Bebortta, S., Barik, S. C., Sahoo, L. K., Mohapatra, S. S., Kaiwartya, O., and Senapati, D., "Hybrid Machine Learning Framework for Network Intrusion Detection in IoT-Based Environments," *Lecture Notes in Networks and Systems*, vol. 1, pp. 573–585, 2024.

Mhaisen, N., Abdellatif, A. A., Mohamed, A., Erbad, A., and Guizani, M., "Optimal User-Edge Assignment in Hierarchical Federated Learning Based on Statistical Properties and Network Topology Constraints," *IEEE Transactions on Network Science and Engineering*, vol. 9, pp. 55–66, 2022.

Ren, X., Wang, Y., Zhang, J., and Han, Z., "Research on Edge-Cloud Collaborative Data Sharing Method Based on Federated Learning in Internet of Vehicles," in *Proceedings of the 2023 IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS)*, Ocean Flower Island, China, Dec. 17–21, 2023, pp. 1075–1080.

Yu, S., Chen, X., Zhou, Z., Gong, X., and Wu, D., "When Deep Reinforcement Learning Meets Federated Learning: Intelligent Multitimescale Resource Management for Multiaccess Edge Computing in 5G Ultradense Networks," *IEEE Internet of Things Journal*, vol. 8, pp. 2238–2251, 2021.

Wang, Z., Xu, H., Liu, J., Huang, H., Qiao, C., and Zhao, Y., "Resource-Efficient Federated Learning with Hierarchical Aggregation in Edge Computing," in *Proceedings of the IEEE INFOCOM 2021—IEEE Conference on Computer Communications*, Vancouver, BC, Canada, May 10–13, 2021, pp. 1–10.

Wu, W., He, L., Lin, W., and Mao, R., "Accelerating Federated Learning Over Reliability-Agnostic Clients in Mobile Edge Computing Systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, pp. 1539–1551, 2021.

Li, Z., He, Y., Yu, H., Kang, J., Li, X., Xu, Z., and Niyato, D., "Data Heterogeneity-Robust Federated Learning via Group Client Selection in Industrial IoT," *IEEE Internet of Things Journal*, vol. 9, pp. 17844–17857, 2022.