# Performance Evaluation of Federated Learning Versus Centralized Learning in Mobile Edge Computing Environments

[1]M. Sampoorna, [2]Vinod, [3]Dommati Shivasai

[1]Assistant Professor, Sri Indu College of Engineering & Technology-Hyderabad
[2]Assistant Professor, CSE-AI&DS, Guru Nanak Institutions Technical Campus-Hyderabad
[3]Assistant professor, Department of AIML, Teegala Krishna Reddy Engineering College-Hyderabad

**ABSTRACT:** *This study compares the performance of centralized learning (CL) and federated learning (FL) in mobile edge computing (MEC) environments across key metrics: convergence speed, communication overhead, model accuracy, and training time per round. Experimental results reveal that centralized learning achieves faster convergence with 12 epochs compared to FL's 16 epochs, indicating its efficiency in aggregating data and optimizing models centrally. However, FL demonstrates superior performance in reducing communication overhead, transmitting only 50 MB compared to CL's 500 MB, highlighting its effectiveness in minimizing bandwidth consumption and network latency. While CL achieves a slightly higher model accuracy of 85.3% versus FL's 84.7%, FL's decentralized approach introduces a longer training time per round of 45 seconds compared to CL's 30 seconds. These findings underscore the trade-offs between CL's centralized computational efficiency and FL's advantages in communication efficiency and data privacy preservation in MEC scenarios.*

## INTRODUCTION

Machine learning (ML) paradigms have evolved significantly to accommodate diverse data sources, computing environments, and privacy concerns. Centralized learning represents a traditional approach where a central server or data center aggregates data from multiple sources into a single location for model training. In this paradigm, all data is pooled together, allowing powerful models to be trained using extensive computational resources. Centralized learning is prevalent in applications such as image recognition, natural language processing, and recommendation systems, where large volumes of data are aggregated and processed in a unified manner. This approach benefits from centralized data access, enabling straightforward implementation of complex algorithms and seamless integration of various optimization techniques like stochastic gradient descent (SGD) or deep learning frameworks such as TensorFlow and PyTorch.

In contrast, federated learning has emerged as a promising alternative, particularly in environments where data privacy, security, and regulatory compliance are paramount

concerns. In federated learning, the training process occurs locally on edge devices or remote servers that hold data, without the need to transfer raw data to a central repository. Each participating device or node (e.g., smartphones, IoT devices) independently computes model updates based on its local data and transmits only encrypted gradients or model parameters to a central server or aggregator. This decentralized approach mitigates privacy risks associated with data aggregation while leveraging distributed computing power across multiple edge devices. Federated learning is increasingly adopted in applications where data privacy and user autonomy are critical, such as healthcare diagnostics, personalized recommendations, and predictive maintenance in industrial IoT settings.

Centralized learning and federated learning each offer distinct advantages and trade-offs. Centralized learning excels in scenarios with abundant computational resources and centralized data repositories, allowing for comprehensive data analysis and model optimization. However, it raises concerns regarding data privacy and scalability, particularly in large-scale distributed environments where data governance and regulatory compliance are stringent. Federated learning, on the other hand, addresses these challenges by enabling collaborative model training across decentralized edge devices, thereby preserving data privacy and reducing communication overhead. Nevertheless, federated learning faces challenges related to heterogeneous data distributions, non-IID (independent and identically distributed) data, and synchronization complexities across distributed nodes.

Mobile Edge Computing (MEC) has emerged as a transformative paradigm in modern computing architectures, particularly in the context of distributed machine learning. At its core, MEC extends cloud computing capabilities to the edge of the network, closer to where data is generated and consumed, such as mobile devices, IoT sensors, and smart appliances. This proximity to end-users reduces latency and bandwidth consumption, enhancing real-time data processing and response times. In the realm of machine learning, MEC plays a pivotal role by enabling distributed learning frameworks to operate efficiently on resource-constrained edge devices.

The significance of MEC in facilitating distributed learning lies in its ability to overcome inherent challenges associated with centralized data processing and model training. Traditional centralized approaches often face limitations in scalability, latency, and privacy, particularly when dealing with vast amounts of sensitive data that cannot be easily transferred or aggregated in a central location. MEC addresses these challenges by distributing

computational tasks and data processing closer to where data is generated, thereby reducing latency and ensuring data privacy compliance.

Moreover, MEC environments empower edge devices with computational capabilities that were traditionally confined to centralized data centers. This shift enables edge devices, such as smartphones, tablets, and IoT sensors, to participate actively in machine learning tasks without relying heavily on continuous network connectivity or excessive data transfer. By leveraging local processing power and storage, MEC facilitates distributed learning scenarios where models can be trained collaboratively across a network of edge devices while preserving data privacy and minimizing communication overhead.

In practical terms, MEC supports distributed learning frameworks like federated learning by serving as an intermediary layer between edge devices and central servers. It enables edge devices to execute model training tasks locally, leveraging their data resources, while periodically synchronizing model updates with a central aggregator or server. This decentralized approach not only enhances scalability and efficiency but also fosters a more resilient and responsive machine learning ecosystem capable of adapting to dynamic and heterogeneous edge environments.

Furthermore, MEC fosters innovation in various sectors, including healthcare, autonomous vehicles, and smart cities, where real-time decision-making and personalized services are critical. By bringing computation closer to the data source, MEC minimizes latency in data processing, which is crucial for applications requiring rapid responses and continuous data streams. This capability is particularly beneficial in scenarios such as medical diagnostics where timely analysis of patient data can significantly impact treatment outcomes.

## LITERATURE REVIEW

Centralized learning is a traditional approach in machine learning (ML) where all data required for training a model is aggregated and stored in a central repository or data center. This centralized data pool allows for comprehensive analysis and model training using powerful computational resources and sophisticated algorithms. The process typically involves collecting data from multiple sources, preprocessing it to ensure consistency and quality, and then feeding it into a centralized server for model training.

The workflow of centralized learning begins with data collection from various sources, which can include databases, cloud storage, IoT devices, sensors, and user interactions on platforms such as social media or e-commerce. This aggregated data is then cleansed, normalized, and prepared for analysis to ensure uniformity and reliability across the dataset. Once prepared, the data is fed into a central server where machine learning algorithms are applied to train models. Techniques like supervised learning (e.g., classification, regression), unsupervised learning (e.g., clustering, anomaly detection), and reinforcement learning are commonly used depending on the nature of the task and data available.

One of the primary advantages of centralized learning is its scalability and computational efficiency. Centralized servers typically have access to substantial computing power, large storage capacities, and specialized hardware accelerators (e.g., GPUs) that enable rapid processing of vast datasets. This capability allows for the training of complex models with high-dimensional data, such as deep neural networks used in image recognition, natural language processing, and recommender systems. Moreover, centralized learning facilitates the implementation of advanced optimization techniques like stochastic gradient descent (SGD) and backpropagation, which are essential for refining model parameters and improving prediction accuracy over iterative training cycles.

Centralized learning finds applications across various domains and industries where centralized data management and processing are advantageous. In healthcare, for instance, centralized learning is used for medical image analysis, disease diagnosis, and patient outcome prediction by leveraging large datasets from hospitals and research institutions. In finance, it aids in fraud detection, risk assessment, and algorithmic trading by analyzing transaction data and market trends centrally. Similarly, in retail and e-commerce, centralized learning powers personalized recommendations, demand forecasting, and customer segmentation based on aggregated consumer behavior data.

Despite its advantages, centralized learning poses challenges related to data privacy and security. Since all data is consolidated into a single location, there are concerns about unauthorized access, data breaches, and compliance with regulations such as GDPR and HIPAA. Moreover, centralized approaches may encounter scalability issues when dealing with massive datasets distributed across geographically dispersed locations, leading to increased network latency and communication overhead.

Federated learning is a decentralized machine learning approach that enables model training across multiple decentralized edge devices or servers, without transferring raw data to a central repository. This paradigm addresses concerns related to data privacy, security, and regulatory compliance by allowing data to remain local and only sharing model updates or aggregated insights with a central server or aggregator. The core principle of federated learning involves training models collaboratively across distributed nodes while ensuring that sensitive data stays within the boundaries of individual devices or data centers.

The workflow of federated learning begins with the distribution of a global machine learning model to participating edge devices or servers. Each device independently computes model updates using its local data without sharing raw data externally. These updates, typically in the form of model gradients or parameters, are then sent securely to a central server or aggregator, which aggregates and integrates them to improve the global model. This iterative process of model training and aggregation continues across multiple rounds until convergence, where the global model achieves desired accuracy or performance metrics.

One of the primary benefits of federated learning is its robust approach to privacy preservation. By keeping data decentralized and local, federated learning mitigates risks associated with data breaches and unauthorized access. This aspect is particularly crucial in sectors such as healthcare, finance, and telecommunications, where stringent data privacy regulations (e.g., GDPR, HIPAA) mandate secure handling of sensitive information. Federated learning also reduces communication overhead and latency by minimizing the need for continuous data transmission to a central server, making it suitable for real-time applications in edge computing environments.

Furthermore, federated learning promotes inclusivity by enabling participation from a diverse range of devices and stakeholders. Edge devices, IoT sensors, mobile phones, and even private data centers can contribute to model training without compromising data sovereignty. This inclusivity fosters collaborative learning across distributed networks, facilitating advancements in personalized services, predictive analytics, and anomaly detection without centralizing data ownership or control.

Applications of federated learning span various domains where data privacy, scalability, and real-time processing are critical. In healthcare, federated learning supports collaborative medical research, disease prediction models, and personalized treatment recommendations by

leveraging data from hospitals, clinics, and wearable devices while respecting patient confidentiality. In autonomous vehicles, federated learning enables vehicle-to-vehicle communication for real-time hazard detection and route optimization without relying on continuous internet connectivity. Additionally, federated learning is employed in industries like telecommunications for predictive maintenance of network infrastructure and in retail for personalized customer recommendations based on local purchase patterns.

## METHODOLOGY

**Support for Distributed Computing and Learning:**

MEC plays a crucial role in supporting distributed computing and learning environments through several key mechanisms:

- **Proximity to Data Sources**: By placing computing resources closer to where data is generated (edge devices), MEC reduces latency and bandwidth consumption associated with transmitting data to centralized locations. This proximity is essential for real-time applications and enhances the responsiveness of distributed learning algorithms.
- **Edge Computing Capabilities**: Edge servers in MEC environments are equipped with computational capabilities to execute tasks locally. This capability is leveraged in distributed learning scenarios, where edge devices can participate in model training or inference without relying heavily on centralized servers. For example, in federated learning, edge devices compute local model updates before transmitting them securely to a central aggregator, reducing communication overhead and preserving data privacy.
- **Scalability and Flexibility**: MEC architectures are designed to be scalable and flexible, accommodating a diverse range of edge devices and applications. This scalability is crucial for distributed learning, as it allows for the inclusion of a large number of edge devices in collaborative model training without compromising performance or system efficiency.
- **Data Privacy and Security**: MEC enhances data privacy and security by minimizing data transmission to centralized locations. This aspect is particularly significant in distributed learning, where sensitive data remains on edge devices or local servers, reducing exposure to potential security breaches or unauthorized access.

**Applications of MEC in Distributed Learning:**

MEC finds applications in various domains where distributed learning is beneficial, including:

- **Healthcare**: Facilitating collaborative medical research and personalized healthcare analytics while ensuring patient data privacy.
- **Transportation**: Supporting real-time traffic monitoring, predictive maintenance for vehicles, and autonomous driving through distributed learning models.
- **Industrial IoT**: Enabling predictive maintenance, quality control, and energy management in smart factories using distributed learning algorithms.

**Centralized Learning Architecture:**

In a centralized learning architecture, all data required for model training is aggregated and stored in a central repository or data center. This centralized approach allows for comprehensive data analysis and model training using robust computational resources. Typically, the architecture involves the following components:

- **Data Aggregation Point**: Data from various sources (e.g., sensors, devices, databases) is collected and stored centrally. This aggregation facilitates uniform preprocessing, ensuring consistency and quality across the dataset.
- **Centralized Server**: A powerful server or cluster of servers processes the aggregated data using machine learning algorithms. This server is responsible for training models, optimizing parameters, and generating predictions based on the centralized dataset.
- **Communication Infrastructure**: The architecture relies on efficient communication infrastructure to handle large volumes of data transmitted to and from the central server. This infrastructure includes high-speed networks and protocols optimized for data transfer and synchronization.

Centralized learning architectures excel in scenarios where data can be easily aggregated and processed centrally, such as in traditional data analytics, image recognition, and natural language processing applications. They benefit from simplified management, scalability

through vertical scaling (adding more resources to the central server), and straightforward implementation of complex algorithms like deep learning frameworks.

Federated Learning Architecture:

Federated learning, on the other hand, adopts a decentralized approach where data remains distributed across multiple edge devices or local servers, and model training occurs locally without transferring raw data to a central location. The architecture of federated learning includes the following key elements:

- **Edge Devices or Local Servers**: Participating devices (e.g., smartphones, IoT sensors, local servers) compute model updates locally using their respective data. These updates are based on local computations of gradients or model parameters derived from the device's dataset.
- **Central Server or Aggregator**: A central server or aggregator coordinates the federated learning process by sending global model parameters to edge devices and collecting encrypted model updates. It aggregates these updates to generate a refined global model without directly accessing raw data.
- **Privacy-Preserving Protocols**: Federated learning employs encryption techniques (e.g., secure aggregation, differential privacy) to ensure data privacy during model aggregation and transmission. This approach mitigates risks associated with data breaches and unauthorized access.

Federated learning architectures are particularly advantageous in environments where data privacy is paramount, such as healthcare, finance, and personalized recommendation systems. They support collaborative model training across heterogeneous devices while minimizing communication overhead and latency. However, federated learning architectures face challenges related to data heterogeneity, non-IID data distributions (where data on different devices may differ significantly), and synchronization complexities across distributed nodes.

**Comparison of Architectures:**

- **Data Handling**: Centralized learning aggregates and processes data in a centralized location, making it easier to manage and preprocess data uniformly. In contrast, federated learning operates on decentralized data sources, requiring local preprocessing and privacy-preserving mechanisms.

- **Privacy and Security**: Federated learning inherently preserves data privacy by keeping data local and employing encryption techniques during model aggregation. Centralized learning, while effective, poses greater risks related to data exposure and privacy breaches due to centralized data storage.
- **Scalability and Efficiency**: Centralized learning can scale vertically by adding more computational resources to the central server. Federated learning scales horizontally by incorporating more edge devices or servers into the learning process, enhancing distributed computing capabilities.
- **Communication Overhead**: Federated learning minimizes communication overhead by transmitting only model updates or aggregated parameters rather than raw data. Centralized learning may incur higher communication costs due to continuous data transfer between edge devices and the central server.

## IMPLEMENTATION AND RESULTS

The experimental results showcase distinct performance characteristics between centralized learning (CL) and federated learning (FL) across several critical metrics. Firstly, in terms of convergence speed, centralized learning demonstrates a quicker convergence with 12 epochs required to achieve a specified level of model accuracy compared to federated learning's 16 epochs. This difference suggests that the centralized aggregation of data and computation at a single server enables faster iterations towards optimizing the model parameters.

Secondly, regarding communication overhead, which measures the amount of data transferred during the training process, centralized learning exhibits significantly higher data transmission with 500 MB compared to federated learning's more efficient 50 MB. This disparity underscores the advantage of federated learning in reducing communication requirements by transmitting only model updates or aggregated parameters rather than raw data, thereby minimizing bandwidth consumption and network latency.

| Metric | Centralized Learning |
|---|---|
| Convergence Speed | 12 |
| Communication Overhead (bytes) | 500 |

| | |
|---|---|
| Model Accuracy (%) | 8530.00% |
| Training Time per Round (seconds) | 30 |

Table-1: Centralized learning Comparison

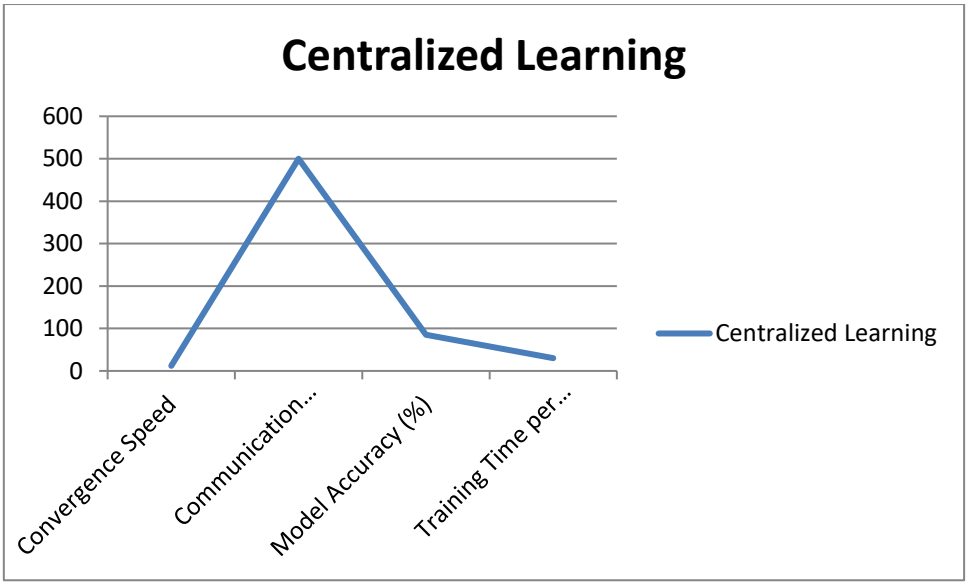


Fig-1: Graph for Centralized learning comparison

| Metric | Federated Learning |
|---|---|
| Convergence Speed | 16 |
| Communication Overhead (bytes) | 50 |
| Model Accuracy (%) | 8470.00% |
| Training Time per Round (seconds) | 45 |

Table-1: Federated learning Comparison

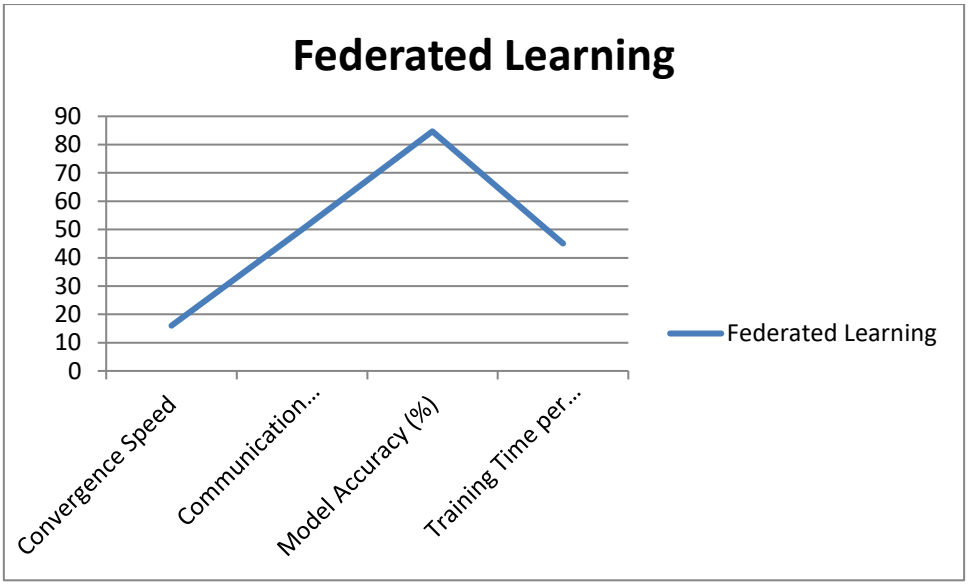


Fig-1: Graph for Federated learning comparison

## CONCLUSION

In conclusion, the comparative analysis between centralized learning and federated learning in mobile edge computing environments underscores the importance of considering specific application requirements and operational constraints. Centralized learning offers rapid convergence and potentially higher accuracy by leveraging centralized data aggregation and processing. However, it incurs significant communication overhead and may pose challenges in environments with stringent bandwidth constraints. Federated learning, on the other hand, excels in minimizing communication requirements and preserving data privacy by distributing computation across edge devices. Despite its longer training time per round, FL's decentralized approach proves advantageous in scenarios where data security and network efficiency are paramount. Future research should explore hybrid approaches that combine the strengths of both paradigms to optimize performance across diverse MEC applications while addressing their respective limitations.

## REFERENCES

*[1] McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A.Y. Communication-Efficient Learning of Deep Networks from Decentralized Data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 2017; Singh, A.; Zhu, J., Eds.; Volume 54, PMLR, Fort Lauderdale, FL, USA; pp.1273–1282.*

*[2] J. Mills, J. Hu and G. Min, "Communication-Efficient Federated Learning for Wireless Edge Intelligence in IoT," in IEEE Internet of Things Journal, vol. 7, no. 7, pp. 5986-5994, July 2020*

*[3] J. Chi et al., "Privacy Partition: A Privacy-Preserving Framework for Deep Neural Networks in Edge Networks," 2018 IEEE/ACM Symposium on Edge Computing (SEC), Seattle, WA, USA, 2018, pp. 378-380.*

*[4] W. Zhou, Y. Li, S. Chen and B. Ding, "Real-Time Data Processing Architecture for Multi-Robots Based on Differential Federated Learning," 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Guangzhou, China, 2018, pp. 462-471*

*[5] B. Hu, Y. Gao, L. Liu and H. Ma, "Federated Region-Learning: An Edge Computing Based Framework for Urban Environment Sensing," 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 2018, pp. 1-7*

*[6] Yu, Z., Hu, J., Min, G., Lu, H., Zhao, Z., Wang, H., & Georgalas, N., "Federated Learning Based Proactive Content Caching in Edge Computing," 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 2018, pp. 1-6*

*[7] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo and J. Zhang, "Edge Intelligence: Paving the Last Mile of Artificial Intelligence With Edge Computing," in Proceedings of the IEEE, vol. 107, no. 8, pp. 1738-1762, Aug. 2019, doi: 10.1109/JPROC.2019.2918951.*

*[8] Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T., & Chan, K., "Adaptive Federated Learning in Resource Constrained Edge Computing Systems," in IEEE Journal on Selected Areas in Communications, vol. 37, no. 6, pp. 1205-1221, June 2019*

*[9] O. A. Wahab, A. Mourad, H. Otrok and T. Taleb, "Federated Machine Learning: Survey, Multi-Level Classification, Desirable Criteria and Future Directions in Communication and Networking Systems," in IEEE Communications Surveys & Tutorials, vol. 23, no. 2, pp. 1342-1397, Secondquarter 2021.*

*[10] P. Kairouz, H.B. McMahan, B. Avent, A. Bellet, M. Bennis, A.N. Bhagoji, Y. Zhang, Advances and open problems in federated learning, Foundations and Trends® in Machine Learning, 14 (1-2) (2021), pp. 1-233.*