A Review of Split Learning and Federated Learning: Challenges and Synergies

Amara Okeke[1], Faridah Abdul[2], Oluwaseun Bamidele[3], and Nneka Obi[1,*]

[1]Department of Computer Science, University of Lagos
[2]Department of Electrical Engineering, University of Nairobi
[3]Department of Computer Science, University of Ibadan
[*]Corresponding author: nneka.obi@unilag.edu.ng

Abstract

Split Learning and Federated Learning have emerged as key techniques in the domain of privacy-preserving distributed machine learning. This paper reviews the recent developments in both paradigms, discussing their respective advantages, limitations, and the potential for their integration. We provide an analysis of current research trends, explore challenges in implementation, and suggest future directions for improving these approaches. The review serves as a resource for researchers and practitioners interested in the evolving landscape of distributed machine learning.

Keywords: Split Learning, Federated Learning, Privacy-Preserving Data, Distributed Systems, Machine Learning.

1. Introduction

The rise of distributed machine learning techniques, particularly Split Learning (SL) and Federated Learning (FL), has been driven by the growing need to process data across multiple locations without compromising privacy Thapa et al., 2022. SL and FL offer frameworks that allow model training on decentralized data, each with unique methods for balancing privacy and performance. This paper provides a comprehensive review of these approaches, examining their current applications, challenges, and the potential for their combined use in privacy-preserving machine learning.

2. Split Learning: Recent Developments

Split Learning (SL) is a distributed learning technique where a deep learning model is split into segments, with different parts trained on different devices. This method significantly

reduces the need to share raw data between devices, making it a powerful tool for privacy-preserving machine learning. Recent developments in SL have focused on improving the efficiency of split points and minimizing the communication overhead between devices Mohammadabadi et al., 2023.

2.1 Applications of Split Learning

SL has found applications in various domains, particularly where data privacy is critical. For example, in healthcare, SL allows the training of models on sensitive patient data without sharing it across institutions Thapa et al., 2021. In finance, SL enables the use of customer data across banks to improve fraud detection while maintaining confidentiality.

2.2 Challenges in Split Learning

Despite its advantages, SL faces challenges such as high communication costs and potential security vulnerabilities at the split points Mohammadabadi et al., 2024. Research efforts are focused on optimizing the trade-off between privacy and computational efficiency, with innovative approaches being explored to reduce the latency and security risks associated with split points.

3. Federated Learning: An Overview

Federated Learning (FL) has become a cornerstone of distributed machine learning, particularly in scenarios where data cannot be moved from its source. FL allows the training of models across multiple devices by only sharing model updates, thus preserving the privacy of individual datasets Zhao et al., 2023. This section explores the current state of FL, its applications, and the ongoing challenges in its implementation.

3.1 Applications of Federated Learning

FL is widely used in areas such as mobile device personalization, where it enables the training of predictive models without collecting user data on a central server L. Li et al., 2020. In the medical field, FL facilitates collaborative research across institutions by allowing the sharing of model knowledge without exchanging raw patient data Zhang et al., 2022.

3.2 Challenges in Federated Learning

One of the primary challenges in FL is the issue of communication efficiency. The process of aggregating model updates from numerous devices can be resource-intensive and slow Mammen, 2021. Moreover, ensuring the robustness and security of the FL process,

particularly against adversarial attacks, remains a critical area of research T. Li et al., 2020.

4. Synergies Between Split Learning and Federated Learning

The integration of Split Learning (SL) and Federated Learning (FL) presents a compelling opportunity to address the inherent limitations of both approaches and create more robust, privacy-preserving distributed learning systems.

Split Learning, which involves partitioning a machine learning model across multiple parties, offers advantages in terms of data privacy and computational efficiency. By distributing model components, sensitive data remains localized, mitigating privacy risks associated with centralized data collection. Additionally, SL can improve computational efficiency by distributing the training workload across multiple devices Singh et al., 2019.

Federated Learning, on the other hand, focuses on training models on decentralized data without sharing raw data. This approach preserves data privacy while enabling collaborative model development. However, FL can suffer from communication overhead and challenges in handling heterogeneous data distributions.

Combining these two techniques holds the potential to create a synergistic approach that leverages the strengths of both while mitigating their weaknesses. For instance, by integrating SL's model partitioning with FL's decentralized training, it is possible to develop distributed learning systems that are both privacy-preserving and computationally efficient. This combination can also enhance the robustness of the system by reducing the impact of data heterogeneity and improving convergence.

Furthermore, the synergy between SL and FL can open up new avenues for research and development. For example, exploring the optimal partitioning strategies for different types of models and data distributions is a critical area of investigation. Additionally, developing efficient communication protocols for exchanging model updates between split model components in a federated setting is essential for practical implementation.

In conclusion, the combination of Split Learning and Federated Learning offers a promising path towards developing more secure, efficient, and scalable distributed learning systems. By carefully considering the strengths and weaknesses of both approaches, researchers and practitioners can unlock the full potential of this synergistic combination.

4.1 Enhanced Privacy and Security

By combining SL and FL, it is possible to enhance data privacy and model security. SL can reduce the amount of information shared during the FL process by splitting the model and only transmitting intermediate results, thus minimizing the risk of data leakage Gao et al., 2020. This approach also allows for more flexible model architectures that can

adapt to different privacy requirements.

4.2 Improving Communication Efficiency

One of the major benefits of integrating SL with FL is the potential to improve communication efficiency. By carefully choosing the split points in SL, it is possible to reduce the amount of data that needs to be exchanged during the FL process Shen et al., 2023. This can lead to faster training times and reduced resource consumption, making the combined approach more suitable for large-scale distributed systems.

Table 1: Comparison of Split Learning and Federated Learning Techniques

Technique	**Model Type**	**Applications**	**Challenges**
Split Learning	Distributed	Privacy-preserving model training	Communication overhead, split point security
Federated Learning	Distributed	Decentralized learning, edge computing	Communication efficiency, robustness
Secure Split Learning	Distributed	Healthcare, finance	Balancing privacy with efficiency
SplitFed Learning	Hybrid	Cross-domain model training	Complexity, implementation challenges

5. Conclusion

Split Learning and Federated Learning represent two of the most promising approaches in the field of privacy-preserving distributed machine learning. While each technique has its own set of challenges, their integration holds significant potential for enhancing privacy, security, and communication efficiency in distributed learning systems. Future research should focus on developing hybrid approaches that leverage the strengths of both SL and FL, ultimately leading to more robust and scalable solutions for real-world applications.

References

Gao, Y., Kim, M., Abuadbba, S., Kim, Y., Thapa, C., Kim, K., Camtepe, S. A., Kim, H., & Nepal, S. (2020). End-to-end evaluation of federated learning and split learning for internet of things. *arXiv preprint arXiv:2003.13376*.

Li, L., Fan, Y., Tse, M., & Lin, K.-Y. (2020). A review of applications in federated learning. *Computers & Industrial Engineering*, 149, 106854.

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE signal processing magazine*, 37(3), 50–60.

Mammen, P. M. (2021). Federated learning: Opportunities and challenges. *arXiv preprint arXiv:2101.05428*.

Mohammadabadi, S. M. S., Yang, L., Yan, F., & Zhang, J. (2024). Communication-efficient training workload balancing for decentralized multi-agent learning. *arXiv preprint arXiv:2405.00839*.

Mohammadabadi, S. M. S., Zawad, S., Yan, F., & Yang, L. (2023). Speed up federated learning in heterogeneous environment: A dynamic tiering approach. *arXiv preprint arXiv:2312.05642*.

Shen, J., Cheng, N., Wang, X., Lyu, F., Xu, W., Liu, Z., Aldubaikhy, K., & Shen, X. (2023). Ringsfl: An adaptive split federated learning towards taming client heterogeneity. *IEEE Transactions on Mobile Computing*.

Singh, A., Vepakomma, P., Gupta, O., & Raskar, R. (2019). Detailed comparison of communication efficiency of split learning and federated learning. *arXiv preprint arXiv:1909.09145*.

Thapa, C., Arachchige, P. C. M., Camtepe, S., & Sun, L. (2022). Splitfed: When federated learning meets split learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(8), 8485–8493.

Thapa, C., Chamikara, M. A. P., & Camtepe, S. A. (2021). Advancements of federated learning towards privacy preservation: From federated learning to split learning. *Federated Learning Systems: Towards Next-Generation AI*, 79–109.

Zhang, T., Gao, L., He, C., Zhang, M., Krishnamachari, B., & Avestimehr, A. S. (2022). Federated learning for the internet of things: Applications, challenges, and opportunities. *IEEE Internet of Things Magazine*, 5(1), 24–29.

Zhao, Z., Mao, Y., Liu, Y., Song, L., Ouyang, Y., Chen, X., & Ding, W. (2023). Towards efficient communications in federated learning: A contemporary survey. *Journal of the Franklin Institute*, 360(12), 8669–8703.