Extending APPFL: Supporting Vertical, Hierarchical, and Decentralized FL for Science

Argonne NATIONAL LABORATORY

Scientific Achievement

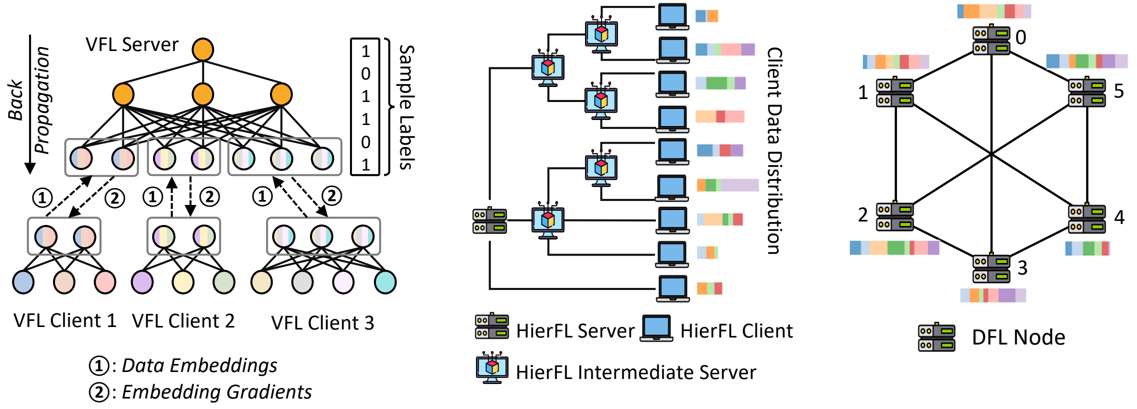
- Extended APPFL to support **vertical**, **hierarchical**, and **decentralized** federated learning (FL), reflecting real-world collaboration patterns across science domains.
- Implemented modular, reusable components for each FL paradigm, enabling flexible and privacy-aware model training across distributed institutions.
- Validated each approach through targeted case studies and benchmarking under heterogeneous conditions.

Significance and Impact

- Broadens the applicability of FL in DOE science use cases where data is siloed, vertically partitioned, or cannot be centrally aggregated.
- Strengthens DOE's software infrastructure for scalable, secure, and distributed AI under the AI for Science (AI4S) initiative.
- Establishes APPFL as a comprehensive, open-source platform for developing and deploying advanced FL algorithms in scientific environments.

Technical Approach

- Built new architecture layers in APPFL to support VFL (split features), HierFL (multi-tier aggregation), and DFL (serverless peer-to-peer learning).
- Enabled secure, cross-site training using Globus IAM and communication layers like Globus Compute and gRPC.
- Evaluated performance across models, data partitions, and system configurations to ensure extensibility and robustness.



The figure illustrates three advanced federated learning paradigms supported by the extended APPFL framework.

- ***VFL** (left) enables collaboration across institutions that hold different features of the same data without sharing raw inputs.*
- ***HierFL** (middle) introduces intermediate servers for multi-level aggregation, reducing communication and reflecting DOE's tiered computing environments.*
- ***DFL** (right) supports peer-to-peer training without a central server, ideal for decentralized, privacy-sensitive settings.*

These extensions make APPFL adaptable to real-world scientific collaboration patterns across DOE and beyond.

PI(s)/Facility Lead(s): Kibaek Kim
Collaborating Institutions: UIUC, Arizona State University
ASCR Program: EXPRESS, AI for Science
ASCR PM: Steven Lee (for both programs)
Publication(s) for this work: Li et al., "Advances in APPFL: A comprehensive and extensible federated learning framework," *IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, 2025 (to appear).