Research Project/Master

SS2024

Towards Convergent Distributed Learning: A Comparative Analysis of Split and Federated Learning in Resource-Constrained Distributed Systems

Project type: ✓ Software ▒ Hardware/Software ✓ Simulation

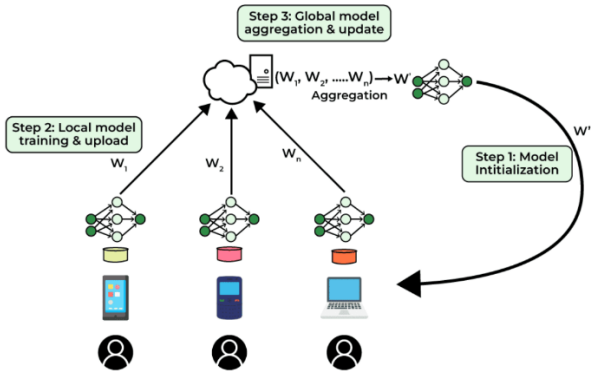
Project description:

Optimization of Wireless Power Transfer Systems Using Meta-heuristics / Genetic Algorithms for Multiobjective Optimization

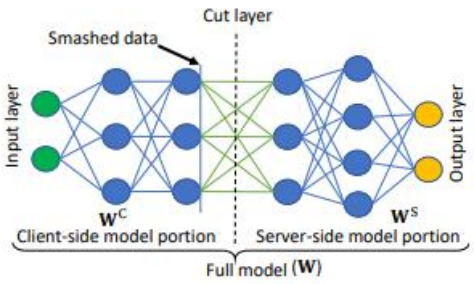
The growth of data generation at the edge, coupled with increasing privacy concerns and resource constraints, necessitates innovative approaches to machine learning (ML) implementation . In response to these challenges, distributed collaborative machine learning (DCML) techniques have emerged, with split learning (SL) and federated learning (FL) as two prominent paradigms. *These techniques allow model training across distributed devices without sharing raw data, instead transferring the ML model to the data for processing.*

However, FL faces challenges in resource-constrained environments due to the computational requirements of training full models on client devices. In contrast, SL, while efficient on single devices, suffers from training time overhead due to its relay-based training across multiple clients. This project aims to perform a rigorous comparative analysis of SL and FL, examining their performance, communication efficiency, and robustness under different conditions focusing on their applicability in resource-constrained environments. Furthermore, recognizing the complementary nature of these techniques, the project will explore the development of hybrid approaches that fuse SL and FL to overcome their limitations and their combined potential for efficient and private distributed learning.

Keywords: Split learning, federated learning, distributed machine learning, comparative analysis, hybrid approach, communication efficiency, resource constraints



FL with N clients



SL with simple setup

Tasks:

1. Literature review: Conduct a comprehensive review of the existing SL and FL literature, identification of research gaps and open challenges

2. Design a flexible and extensible experimental framework to evaluate SL and FL under various conditions, incorporating diverse datasets, model architectures, and a simulation environment for distributed settings
3. Select and preprocess datasets with different distributions and characteristics
4. Implement and optimize SL and FL algorithms for the chosen model architectures
5. Perform a comprehensive comparative analysis of SL and FL along the dimensions of privacy, communication efficiency, scalability, and robustness
6. Investigate and develop novel hybrid SL-FL approaches based on the findings of the comparative analysis
7. Evaluate the performance of hybrid approaches in simulated resource-constrained environments and compare them to single techniques and state-of-the-art met

Competences:

- Distributed Systems Knowledge: Understanding of distributed system architectures, communication protocols, and network topologies. Familiarity with data/model parallelism and distributed optimization.
- Programming and software engineering: Knowledge of Python and ML libraries (PyTorch, TensorFlow Federated).
- Creative thinking and Problem-Solving Skills and scientific curiosity

Contact:

Dipl. -Ing. Oumayma Kahouli
Chair of Measurement and Sensor Technology
Reichenhainerstr. 70 Weinholdbau W283
Email: oumayma.kahouli@etit.tu-chemnitz.de