Federated Transfer Learning

Johannes Appelt, Andreas Herzinger

Transfer Learning

- **Core Idea**: Knowledge learned in one domain can be applied to a different but related domain
- Reduces the amount of data and computational resources needed for training
- Powerful when there is limited training data in target domain

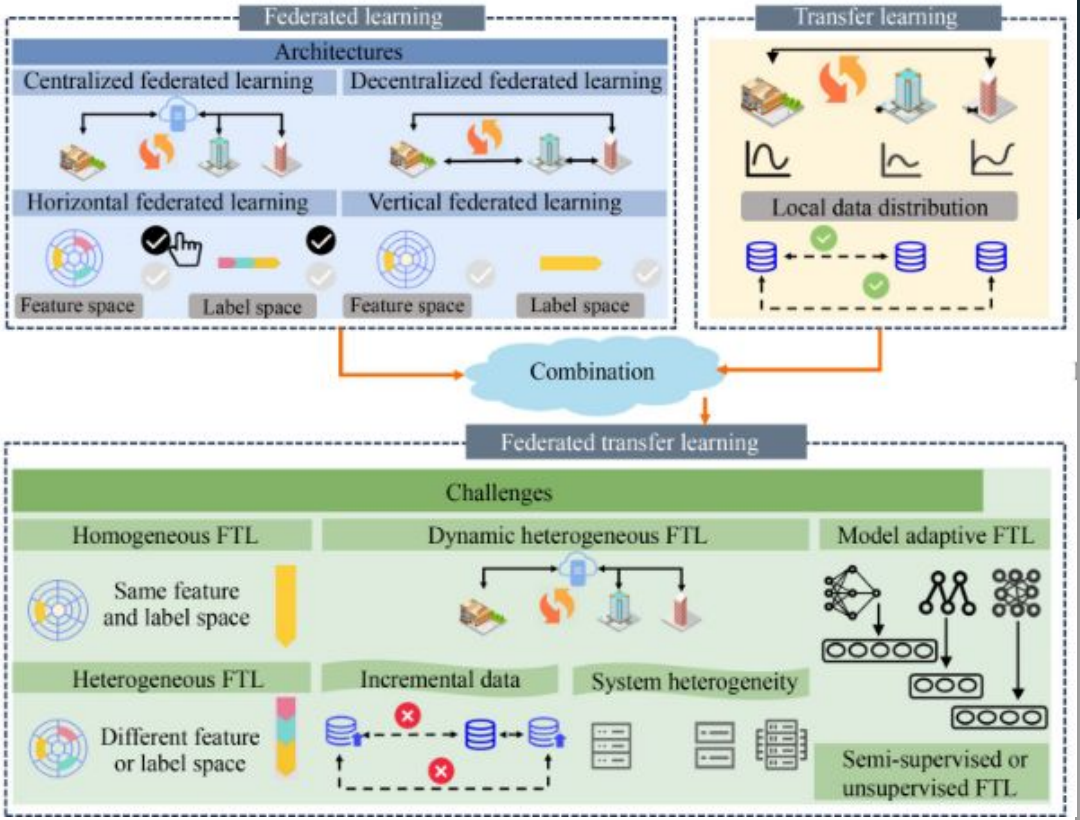
Feature Extraction: Get pre-trained model's parameters -> Replace final classification layer -> Train only last layer

Fine-Tuning: Adjust pre-trained model weights on new dataset

Definition of FTL

FTL = Federated Learning + Transfer Learning

Transfer Learning: transfer pretrained model to different domain



source: https://doi.org/10.1007/s11704-024-40065-x

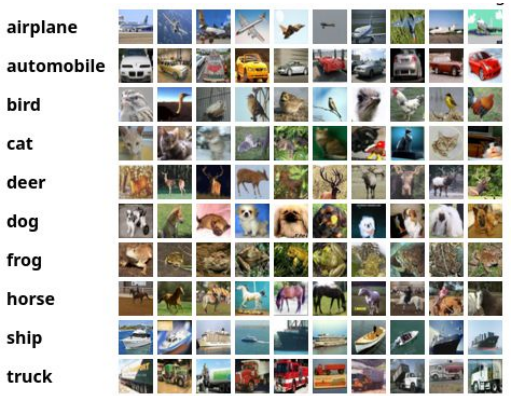
Homogeneous/Heterogeneous FTL

Homogeneous -> IID Client Data, Heterogeneous -> Non-IID Client Data

main Idea: use same pre-trained model for identical FL-setup and only differ Data Distribution and FL-Strategy

- use CIFAR-10 dataset
- split one normally, for other notebook create Non-IID Data artificially (FedArtML)
- use PyTorch for Transfer Learning (resnet18)
- use Flower-Framework for FL

-> Goal: compare performance of pre-trained, homogeneous and heterogeneous



Practical Application Example

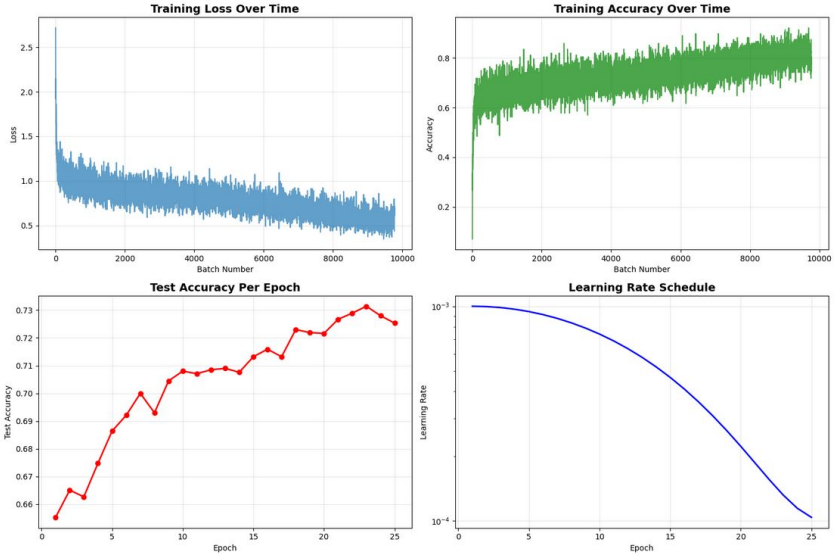
Computer Vision in Medical Images:

- Hospitals usually have only little amount of data -> practically impossible to train model only for one hospital
- a lot of the data may be Non-IID (different specializations of hospitals)
- data privacy is really important (sensitive patient data)

--> use FTL with pretrained model to fine-tune using data of multiple hospitals

Results Federated Transfer learning with PyTorch and Flower using cifar-10 and ResNet-18

- Evaluated FTL setup using IID (homogeneous) client data partitioning
- Consistent class distribution across clients led to stable and predictable training
- Achieved strong convergence and high accuracy with fewer communication rounds
- Pre-trained model adaptation (transfer learning) further improved performance
- FedAvg strategy worked efficiently, requiring minimal tuning for homogeneous data



Results Non-IID Data

- Tested FTL setup with multiple Partitioners and Strategies
- FTL performed better than expected, especially using Dirichlet-Partitioning
- Even for low (Dirichlet) alpha the models learned well
- FedAvg serves as good default, using a different one requires a lot of adaptation and caution