Development and evaluation of global aggregation methods in federated learning

Background

Federated learning (FL) promises to revolutionize medical AI by enabling collaborative training without sharing sensitive data (see Figure 1). However, current FL methods fail catastrophically with real-world heterogeneous medical data, involving different scanners, protocols, populations, and inconsistent clinical annotations, which cause model divergence and poor performance. This fundamental limitation prevents millions of medical images (and other medical data) from being used for AI development. At the same time, a critical shortage of radiologists and pathologists creates an urgent healthcare crisis, which can be solved using AI models in healthcare. To overcome challenges with heterogeneous and inconsistent data is crucial for taking full advantage of FL, and this master thesis will therefore evaluate and develop novel global aggregation functions.

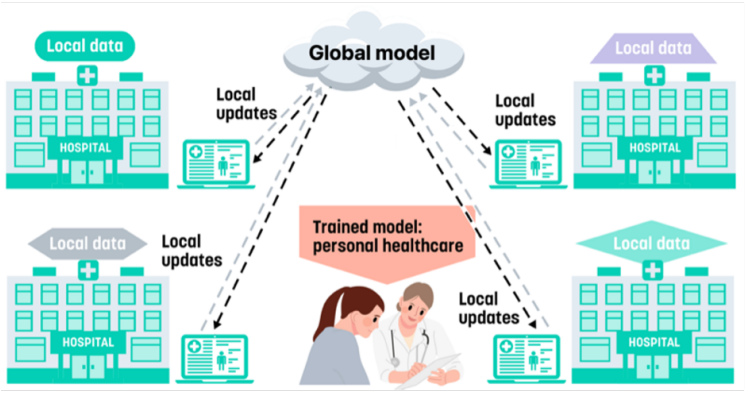

Figure 1: *Federated learning solves the data sharing problem, such that sensitive data from several nodes can be combined for training of deep models, without sending any data between the nodes. Instead, the deep model is in real-time communicated between each node and the global server (the cloud in this image) during training.*

Federated averaging (FedAvg) was proposed in the seminal paper of McMahan et al. (2017), and is still one of the most common FL algorithms. FedAvg provides a global model (for all nodes) by aggregating local models (from each node independently), and is guaranteed to converge for homogenous data. For FedAvg with *K* nodes, the global model weights $\mathrm{w_g}$ are aggregated as $\mathrm{w_g} = \sum_{\mathrm{k=1}}^{\mathrm{K}} \alpha_{\mathrm{k}} \mathrm{w_k}$, where $\mathrm{w_k}$ are the model weights at node *k* and $\alpha_k = \frac{n_k}{N}$, with n_k the number of training samples at node *k* and *N* the total number of training samples. In other words, the global model $\mathrm{w_g}$ is in FedAvg a weighted average of the models $\mathrm{w_k}$ at each node *k*, where the weight is the proportion of data at each node. FedAvg is, however, not guaranteed to converge when the nodes have heterogeneous data. To solve this, new algorithms and aggregation methods have been proposed. For instance, FedProx (Li et al., 2020) is a generalisation of FedAvg where each node instead resolves a proximal operator, leading to a regularised node loss, essentially forcing the local models to not deviate too much from the global model. FedProx is guaranteed to converge also with heterogeneous data, but convergence may take a very long time. Other aggregation methods include inverse distance aggregation (IDA) (Yeganeh et al., 2020) and robust aggregation using the geometric median (Pillutla et al., 2022). However, these approaches only consider differences in the raw model weights.

This master thesis will instead develop novel aggregation functions that also consider differences in the local data, hypothesizing that data differences are important but overlooked. Specifically, the aggregation weights will instead be based on how similar the learned neural network representations are. We will achieve this by developing a federated version of centered kernel alignment (CKA, FedCKA) (Kornblith et al., 2019), see Figure 2. As FedCKA is calculated from activations in each layer of a neural network it thereby takes the data at each node into account.

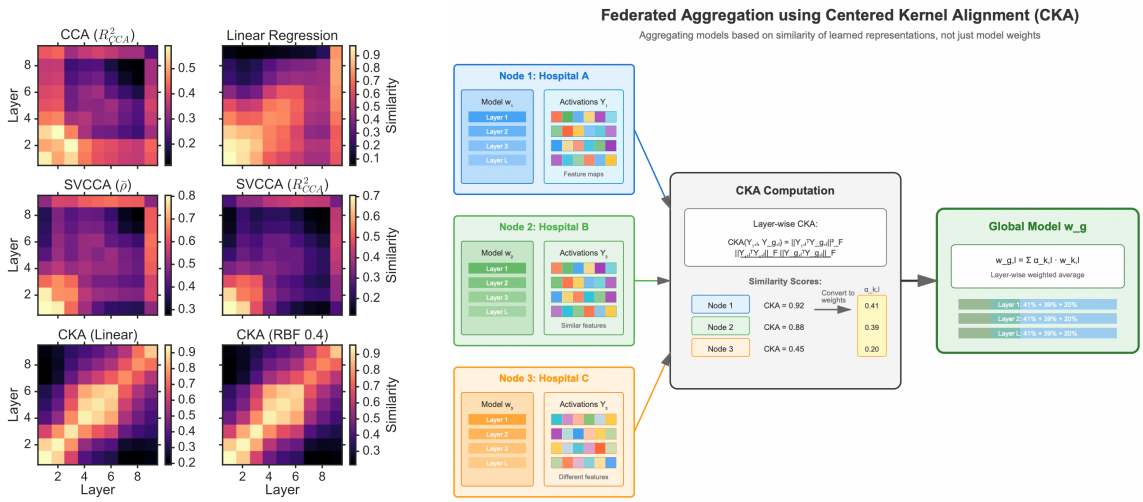

Figure 2. Left: *CKA reveals consistent relationships between layers of CNNs trained with different random initializations, whereas other metrics such as canonical correlation analysis (CCA), singular vector CCA (SVCCA) and linear regression do not (Kornblith et al., 2019). CKA can thereby better determine if two networks have learned a similar representation. This will be used to calculate aggregation weights in the federation, instead of simply comparing raw model weights.* ***Right:*** *Nodes with similar representations will get a higher aggregation weight. The aggregation weights will be calculated separately for each layer, as we expect shallow layers to learn more similar representations.*

Objectives

Perform federated training with different datasets, using different global aggregation functions. Compare the performance of the federated approach to local and centralized training.

Implement federated centered kernel alignment as a new global aggregation approach, and compare it to existing aggregation functions.

Make the different datasets more heterogeneous, to study how the different aggregation functions handle data with different degrees of heterogeneity.

If time permits, extend CKA to distributions of activations, to instead compare distributions of neural network similarities.

Data

Several open datasets will be used, for example containing medical images.

Required background

Machine learning, deep learning, Python programming

Computing resources

The student will have access to very good computing resources (graphics cards).

Contact persons

Anders Eklund, anders.eklund@liu.se, Department of Biomedical Engineering, Department of Computer and Information Science, Center for Medical Image Science and Visualization

References

Simon Kornblith et al. "Similarity of neural network representations revisited". In: International conference on machine learning. PMLR. 2019, pp. 3519–3529.

Tian Li et al. "Federated optimization in heterogeneous networks". In: Proceedings of Machine learning and systems (2020), pp. 429–450

Brendan McMahan et al. "Communication-efficient learning of deep networks from decentralized data". In: Artificial intelligence and statistics. PMLR. 2017, pp. 1273–1282

Krishna Pillutla, Sham M Kakade, and Zaid Harchaoui. "Robust aggregation for federated learning". In: IEEE Transactions on Signal Processing 70 (2022), pp. 1142–1154

Yousef Yeganeh et al. "Inverse distance aggregation for federated learning with non-iid data". In: Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning: Second MICCAI Workshop. Springer. 2020, pp. 150–159