University of Pittsburgh | School of Computing and Information

Proposal Defense

Doctor of Philosophy in Intelligent Systems

"**Personalized Federated Learning over Heterogeneous Data**" by **Jun Luo**

Date: Dec. 11, 2024
Time: 10:00 a.m. – 11:30 a.m.
Place: Room 6106, Sennott Square, 3810 Forbes Ave, Pittsburgh, PA 15213

Committee:

- Dr. Shandong Wu, Intelligent Systems Program, School of Computing and Information
- Dr. Leming Zhou, Intelligent Systems Program, School of Computing and Information
- Dr. Xiaowei Jia, Intelligent Systems Program, School of Computing and Information
- Dr. Lu Tang, Department of Biostatistics, School of Public Health

Abstract:
Training machine learning models for data owners who host privacy-sensitive datasets is challenging as it is often impractical to centralize their data for training. Federated learning (FL) has emerged as a decentralized machine learning technique that offers a collaborative solution to this scenario. In conventional FL, a consensus model is trained by aggregating the locally updated copies of a global model using different private datasets. However, focusing on training a single consensus global model often encounters data heterogeneity issue, i.e., the model often exhibits suboptimal generalization when clients' data is not independent or identically distributed (IID). One way to mitigate this issue is personalized federated learning (PFL), where customized models for different clients are allowed, improving the generalization of the models.

In this thesis proposal, we present a suite of four proposed FL methods for mitigating the data heterogeneity issue, with a major focus on the personalization aspect. Among these methods, the first method, FedSLD, acquires the label distribution from a global perspective, and shares this information to each client for weighting each data sample accordingly. The second method, APPLE, proposes adaptive personalized aggregation, which enables each client to adaptively learn how much they can benefit from other clients' models. It also leverages regularization to control the training between global collaboration and local personalization. Based on this idea, the third method, PGFed, evaluates the downloaded non-local models with local dataset explicitly to learn the adaptive aggregation weights. Additionally, it utilizes a first-order approximation to personalize the aggregation on the server, while asymptotically reducing the communication overhead. Lastly, inspired by these works while leveraging modern Vision-Language Models (VLMs), we propose a novel PFL framework dubbed pFedMoAP. pFedMoAP benefits from VLMs' remarkable capabilities in learning transferable representations across heterogeneous tasks. It also implements the lightweight prompt learning and allows sharing of pre-aggregated prompts, which enables a Mixture of Experts (MoE) perspective to mitigate the data heterogeneity issue.

To evaluate the effectiveness of the proposed methods, we conduct extensive experiments with various datasets under different federated settings. To mimic pathological and practical non-IID settings, we simulate real-world federated settings using techniques including hand- crafted shards of dataset, random selection of labels, and Dirichlet distribution. The experimental results demonstrate the proposed methods' superiority over compared FL/PFL algorithms.