Research Project: Build Edge Based Federated Learning Framework for Personalized Health Monitoring

Introduction & Motivation

With the rapid development of computing technology, wearable devices such as smartphones, wristbands and Smart healthcare IOT devices make it easy to get access to people's health information including activities, sleep, sports, Temperature, BP Level etc. Smart healthcare achieves great success by training machine learning models on large quantities of user data. However, there are two critical challenges. Firstly, user data often exists in the form of isolated islands, making it difficult to perform aggregation without compromising privacy security. Secondly, the models trained on the cloud fail on personalization.

- When the same user uses different products from two companies, his data stored in two clouds cannot be exchanged. This makes it hard to train powerful models using these valuable data.
- Most of the methods are based on a common server model for nearly all users, But different users have different physical characteristics and daily activity patterns. Therefore, the common model fails to perform personalized healthcare.

Existing studies of Federated Learning focus on training a single global model, which would suffer from the weaknesses in both statistical and communication perspectives and lacks personalization, resulting in a degraded performance in the healthcare scenarios. The personally-generated data naturally exhibits the kind of non-IID (non-independent and identically distributed) distribution. Even for a single user, the health monitoring data can be highly skewed. For example, an adult's activity data may include a lot of standing and walking samples, but very few falls. Both the imbalanced and non-IID distribution of user health data may greatly degrade the learning performance. The communication challenge refers to communication constraints, such as slow or expensive connections and limited bandwidth costs. Existing federated learning approaches mainly focus on the communication rounds reduction considering the fact that mobile devices are frequently offline . However, the communication overload reduction in each round is also critical to improve the efficiency of the algorithm considering the bandwidth cost. Finally, as for personalization, the shared model trained by federated learning only captures the common features of all users, but it may perform poorly on a particular user. Thus, it is necessary to learn the fine-grained information on a particular user for personalized healthcare. However, current federated learning researches in healthcare only focus on one of these challenges, without considering them as an interrelated integration.

Aims and objectives

This project aims to develop a novel efficient federated learning framework over decentralized data while preserving privacy. The model will be tested and evaluated with data sets in the wearable healthcare domain.

Objectives:

- Develop FL framework for wearable healthcare, which aggregates the data from different organizations without compromising privacy security.
- Build Personalized model learning with user's local data.

Reference Data Set

1. https://data.world/nasa/wearable-health-monitoring
2. https://www.cs.dartmouth.edu/~dfk/research/rawassizadeh-datasets/rawassizadeh-datasets.pdf
3. https://archive.ics.uci.edu/ml/datasets/human+activity+recognition+using+smartphones
4. https://www.kaggle.com/arashnic/fitbit

References Papers/Resources/Baseline Code

- Federated learning Base Paper https://arxiv.org/pdf/1602.05629.pdf
- https://arxiv.org/pdf/1907.09173.pdf
- https://arxiv.org/pdf/1908.07873.pdf
- https://dl.acm.org/doi/abs/10.1145/3384419.3430446
- https://arxiv.org/abs/2012.07450
- https://arxiv.org/abs/2010.15561
- https://www.sciencedirect.com/science/article/pii/S0926580521001898
- https://arxiv.org/abs/2002.07948
- https://arxiv.org/abs/2003.08673
- https://github.com/chaoyanghe/Awesome-Federated-Learning
- https://flower.dev/
- https://github.com/OpenMined/PySyft
- https://github.com/pliang279/LG-FedAvg