Demo Abstract: Federated Learning on Wearable Devices

Xiaoxin He[1], Xiang Su[2,3], Yang Chen[1], Pan Hui[2,4]
[1]School of Computer Science, Fudan University, China
[2]Department of Computer Science, University of Helsinki, Finland
[3]Center for Ubiquitous Computing, University of Oulu, Finland
[4]Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Hong Kong
{xxhe17,chenyang}@fudan.edu.cn,{xiang.su,pan.hui}@helsinki.fi

ABSTRACT

Wearable devices collect user information about their activities and provide insights to improve their daily lifestyles. Smart health applications have achieved great success by training Machine Learning (ML) models on a large quantity of user data from wearables. However, user privacy and scalability are becoming critical challenges for training ML models in a centralized way. Federated learning (FL) is a novel ML paradigm with the goal of training high quality models while distributing training data over a large number of devices. In this demo, we present FL4W, a FL system with wearable devices enabling training a human activity recognition classifier. We also perform preliminary analytics to investigate the model performance with increasing computation of clients.

CCS CONCEPTS

• **Human-centered computing** → **Mobile computing**.

KEYWORDS

Federated learning, wearable devices, human activity recognition

1 INTRODUCTION

Wearable devices, such as smart watches, fitness trackers, and wristbands, track physical activities of people. Machine learning (ML) algorithms analyze collected data from wearable devices and provide insights to users regarding to physical activities. This paves the way for implementing novel services into populations for managing and optimizing lifestyles. However, privacy remains a key obstacle to implement data analytic algorithms for wearable systems. For example, users may be skeptical in allowing their personal data to be analysed by ML algorithms on Cloud, because of potential privacy risks. Hence, it would be ideal to store the sensitive data only on the trusted devices of users.

Federated learning (FL) [3] is a new ML paradigm, enabling training and inference in a decentralized fashion on heterogeneous devices. Key benefits of FL pertain to 1) the mitigation of systemic privacy risks because clients store their own data; 2) consistently trained and updated model facilitating real-time decision making; and 3) saving network bandwidth as only updated model parameters for aggregation are required to be transmitted to the server machine. In principle, FL fulfills the privacy and scalability requirements for wearable computing with keeping personal data and models on the devices of users and enables them to have direct and physical control of their own data and models [2]. We present FL4W, a FL system for resource-constrained wearable devices. We conduct preliminary experiments of evaluating FL for the task of Human Activity Recognition (HAR) with a daily and sport activities dataset. Our demo shows that 1) FL4W system is capable of producing high accuracy models for the HAR task; and 2) less communication rounds are required to achieve an acceptable accuracy with increasing computation per device. As far as we know, this is one of the first efforts on implementing FL on wearable devices.

2 FL4W: A FL SYSTEM FOR WEARABLES

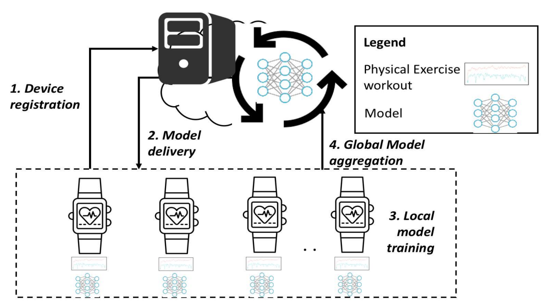

Figure 1: The FL System Architecture on wearable devices.

Figure 1 presents our implemented system architecture of FL4W. The system design follows a client-server architecture. Server machine coordinates participating wearables to train a shallow multilayer perceptron (MLP) model for HAR. One round of collaborative training contains four steps, including device registration, model delivery, local model training, and global model aggregation. Here, the local model refers to the model trained on wearables based on physical exercise workout and the global model refers to the aggregated model at the server machine. **Device registration:** participating wearables can check-in with the FL server, and rejected wearables are informed to participate later on. **Model delivery:** FL server specifies the training task and hyperparameters, and broadcasts the initial model and configurations to participating wearables. **Local model training:** Each local wearable utilizes local data, i.e., physical exercise workout in our scenario, and their own computation resources to update local models. Local training datasets will never be uploaded to the FL server. **Global model aggregation:**

SenSys '20, November 16–19, 2020, Virtual Event, Japan

ACM ISBN 978-1-4503-7590-0/20/11…$15.00
https://doi.org/10.1145/3384419.3430446

The parameter tables of updated local models are subsequently sent to the FL server. The FL server aggregates the local models with Federated Averaging (FedAvg) [3] algorithm. A number of rounds of communication between the participating wearables and the FL server are required to achieve an acceptable accuracy.

3 DEMO SETUP AND PRELIMINARY RESULTS

Demo setup. HAR is a classification ML task targeting the learning of which activity is performed by a certain person in a given period of time. HAR classifies sequences of accelerate data recorded by wearable devices into known well-defined movements. FL enables training a HAR classifier in a privacy-preserving fashion. We setup our demo in a controlled environment with a HAR dataset. The dataset is built from a Daily and Sports Activities Data Set (DSA) [1], which comprises of motion sensor data of ten daily sports activities each performed by eight subjects. A typical FL system contains participating wearables with local datasets, which follow different distributions, i.e., the datasets of participating wearables are non-IID. Non-IID means the training data on a given client is typically based on the usage of the client by a particular user, hence any particular user's local dataset will not be representative of the population distribution [3]. We first sort the data by activity label, dividing into 200 shards of size 2400, and assigning each 100 clients, 2 shards. This is a pathological non-IID partition of the data, as most clients will only have examples of two activities. This DSA dataset is hereby divided into a training and an inference dataset, which are 83.3% and 16.7% of the whole dataset, respectively. The training interactions are repeated over many epochs, i.e., full passes over the training set, until convergence. The selected model is a MLP, composed of one input and one output layer, and one hidden layer with 1000 units using ReLU activations. MLP is often applied to supervised learning problems and it appears to be capable in obtaining a decent performance with a reasonable amount of epochs and 500,000 samples. The volume of data from wearables is small considering the limited storage capability of devices, while data is non-IID as a result of varying user habits. We transfer the learned features from server to clients. In practice, we freeze the the weights of the first two layers of MLP model and retrain the subsequent output layer. Server of FL4W is deployed on a MacBook Pro Laptop equipped with a 2.9GHz Intel Core i5 processor, 16GB memory and an Intel Iris GPU. The client runs on a WearOS simulator and a TicWatch Pro LTE smartwatch. Both the server and client of FL4W are implemented with Java and the source code is available from https://github.com/Cautiousss/federated-learning.

Preliminary results. Inspired by [3], we test accuracy of trained MLP model with increasing computation per client. We report the number of communication rounds necessary to achieve an acceptable accuracy (acc=0.95), when each client performs a more complex calculation. Each wearable device locally takes gradient descent on the local model using local data, then the server takes a weighted average of the resulting model. The amount of computation is controlled by: C, the fraction of clients that perform computation on each round; E, the number of training passes each client makes over its local dataset on each round; and B, the local minibatch size used for the client updates. $B = \infty$ indicates that the full local dataset is treated as a single minibatch. Therefore, we utilize $B = \infty$ and $E = 1$ which corresponds exactly to FedSGD, which is considered as a baseline algorithm. In Figure 2, we fix C = 0.1 and add more computation per client on each round, either decreasing B, increasing E, or both. Figure 2 presents the learning curves of base line and ($B = \infty$, E=5), (B = 50, E=1), (B = 10, E=1), (B = 50, E=5), (B = 10, E=5), respectively. Our results demonstrates that 1) FedAvg converges to a higher level of accuracy than the baseline FedSGD; and 2) adding more local SGD updates per round can significantly decrease communication costs. The expected number of updates per client per round is $\mu = En/KB$, where the expectation is over the draw of a random client k. We see that increasing μ by varying both E and B is effective. It is noted that increasing B is taking advantage of available computation resources of wearable devices and in practice, this should be a primary parameter to be tuned.

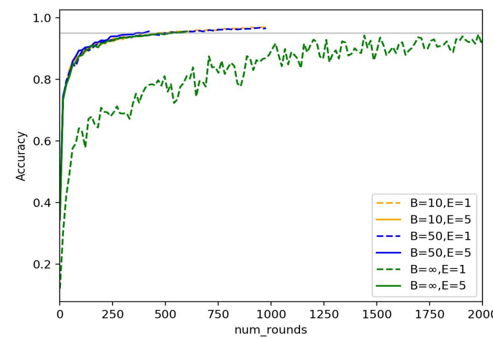

Figure 2: Test set accuracy vs. communication rounds with C = 0.1 and optimized $\eta = 0.05$.

4 CONCLUSION

FL allows for development of smarter models while ensuring privacy. In addition to provide an update to the global model, the improved model on wearables can be used immediately, powering personalized experiences to users. We demo a FL system for wearable devices with focusing on training a HAR classifier. Our system demonstrates the feasibility to develop a FL system with wearables, which paves the way for the development of many potential applications than HAR. Our preliminary results verify that a FL system can achieve a target accuracy. Increasing computation on clients and utilising parallelism enable less communication costs.

ACKNOWLEDGMENTS

This work is sponsored by CERNET Innovation Project (NGII20190105) and Academy of Finland grant 325774 and 325570.

REFERENCES

[1] Billur Barshan and Murat Cihan Yüksek. 2013. Recognizing Daily and Sports Activities in Two Open Source Machine Learning Environments Using Body-Worn Sensor Units. *The Computer Journal* 57, 11 (2013), 1649–1667.

[2] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao. 2020. FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare. *IEEE Intelligent Systems* 35, 4 (2020), 83–93.

[3] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. 2017. Communication-Efficient Learning of Deep Networks from Decentralized Data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics.*