Personal privacy

Physical privacy could be defined as preventing “intrusions into one's physical space or solitude.

An example of the legal basis for the right to physical privacy is the U.S. Fourth Amendment, which guarantees “the right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures”

Physical privacy may be a matter of cultural sensitivity, personal dignity, and/or shyness. There may also be concerns about safety, if for example one is wary of becoming the victim of crime or stalking.

Federated learning (FL) is a newly proposed machine learning method that uses a decentralized dataset. Since data transfer is not necessary for the learning process in FL, there is a significant advantage in protecting personal privacy. Therefore, many studies are being actively conducted in the applications of FL for diverse areas.

Objective: The aim of this study was to evaluate the reliability and performance of FL using three benchmark datasets, including a clinical benchmark dataset.

Methods: To evaluate FL in a realistic setting, we implemented FL using a client-server architecture with Python. The implemented client-server version of the FL software was deployed to Amazon Web Services. Modified National Institute of Standards and Technology (MNIST), Medical Information Mart for Intensive Care-III (MIMIC-III), and electrocardiogram (ECG) datasets were used to evaluate the performance of FL. To test FL in a realistic setting, the MNIST dataset was split into 10 different clients, with one digit for each client. In addition, we conducted four different experiments according to basic, imbalanced, skewed, and a combination of imbalanced and skewed data distributions. We also compared the performance of FL to that of the state-of-the-art method with respect to in-hospital mortality using the MIMIC-III dataset. Likewise, we conducted experiments comparing basic and imbalanced data distributions using MIMIC-III and ECG data.

Results: FL on the basic MNIST dataset with 10 clients achieved an area under the receiver operating characteristic curve (AUROC) of 0.997 and an F1-score of 0.946. The experiment with the imbalanced MNIST dataset achieved an AUROC of 0.995 and an F1-score of 0.921. The experiment with the skewed MNIST dataset achieved an AUROC of 0.992 and an F1-score of 0.905. Finally, the combined imbalanced and skewed experiment achieved an AUROC of 0.990 and an F1-score of 0.891. The basic FL on in-hospital mortality using MIMIC-III data achieved an AUROC of 0.850 and an F1-score of 0.944, while the experiment with the imbalanced MIMIC-III dataset achieved an AUROC of 0.850 and an F1-score of 0.943. For ECG classification, the basic FL achieved an AUROC of 0.938 and an F1-score of 0.807, and the imbalanced ECG dataset achieved an AUROC of 0.943 and an F1-score of 0.807.

FL demonstrated comparative performance on different benchmark datasets. In addition, FL demonstrated reliable performance in cases where the distribution was imbalanced, skewed, and extreme, reflecting the real-life scenario in which data distributions from various hospitals are different. FL can achieve high performance while maintaining privacy protection because there is no requirement to centralize the data [1].

1)

Lee GH, Shin SY. Federated Learning on Clinical Benchmark Data: Performance Assessment. J Med

Internet Res. 2020 Oct 26;22(10):e20891. doi: 10.2196/20891. PMID: 33104011.