

Federated vs. Centralized learning: A Comparative Analysis of CNN-Based Myocardial Infarction Classification

Daniel Vikan, Bjørn-Jostein Singstad

University of South-Eastern Norway
Horten, Norway

Introduction: The era of deep learning has demonstrated that the performance of deep neural networks typically improves as the amount of training data increases. Consequently, collaborative efforts between hospitals, enabling model training on larger and more diverse patient cohorts, will in most cases lead to more accurate and robust models. However, the sensitive nature of the data and privacy issues hinder data sharing across institutions. Federated learning (FL) has been proposed as a solution to this problem, enabling institutions to collaboratively train a shared model without exchanging raw data. We here present a practical FL system for electrocardiogram data using an established FL framework, which we validate on two open datasets to build a joint myocardial infarction (MI) classification model.

Methods: We used the PTB-XL and PTB Diagnostic databases, treating the associated MI diagnosis as a binary outcome. Both datasets were split once into training (80%) and test (20%) sets, and this split was kept throughout all experiments. We then trained an InceptionTime-based convolutional neural network under three scenarios: (1) isolated training, (2) centralized training following data aggregation, and (3) FL. For the FL setup, we implemented an end-to-end Flower-based software for binary classification of 12-lead ECGs. This software enables two or more clients to asynchronously exchange model weights with a central server.

Results and Discussion:

Table 1 summarizes the outcomes of our experiments. While centralized training achieved the best performance on PTB-XL, FL exceeded both

Table 1. AUROC for 5-fold isolated, centralized and federated training

	PTB Diag.	PTB-XL
Isolated training	0.54	0.90
Centralized training	0.67	0.92
Federated learning	0.68	0.78

isolated and centralized training on PTB Diagnostic, and isolated training outperformed FL on PTB-XL. One possible reason PTB Diagnostic gains more from FL than PTB-XL is that PTB-XL is almost ten times larger; however, why FL surpasses centralized training on PTB Diagnostic remains unclear.