

INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER, UK



University of Westminster, Coat of Arms

CrowdFORGE: Federated Learning for Optimizing Crowdsourced Contributions in Centralized Global Models

A dissertation by

Mr. D.V. Sandaru Rashmika

W1867090 | 20210334

Supervised by

Mr. Guhanathan Poravi

Submitted in partial fulfilment of the requirements for the
BEng (Hons) Software Engineering degree at the University of Westminster.

April 2025

ABSTRACT

Increasing utilization of distributed, crowdsourced data in federated learning brings forth challenges of non-IID distributions of data, privacy, and computational limitations on resource-constrained devices. Existing federated learning systems compromise on model accuracy for privacy protection and computational efficiency. In this work, we present CrowdForge, a federated learning system designed to achieve maximum global model accuracy by intelligent aggregation of heterogenous crowdsourced inputs without compromising user privacy.

CrowdForge integrates adaptive aggregation techniques, including FedAvg, FedProx, and AdaptiveFedAvg, along with Differential Privacy and Secure Multi-Party Computation to improve security without compromising performance. Through quality-aware filtering and lightweight feature engineering, only high-value contributions are included in the global model. Experimental results demonstrate that CrowdForge improves federated learning accuracy by an average of 8.2% across non-IID datasets compared to traditional federated methods. Moreover, the framework achieves 35% communication overhead reduction and thereby improves the training efficiency on low-resource devices.

Benchmark experiments confirm that CrowdForge sustains 90% model performance on CIFAR-10 under non-IID partitions, beating conventional federated approaches by 6% while ensuring privacy. The system can effectively cope with resource scarcity by reducing per-device computation by 30% through adaptive learning and selective model updates. These findings confirm that CrowdForge is an efficient, scalable, and privacy-friendly federated learning system that can improve global model accuracy in real-world resource-scarce environments.

Subject Descriptors:

- Computing methodologies → Artificial intelligence → Distributed artificial intelligence → Multi-agent systems
- Computer systems organization → Architectures → Distributed architectures → Peer-to-peer architectures

Keywords: Federated Learning, Crowdsourcing, Model Aggregation, Artificial Intelligence, Deep Learning, Data Science