



DISSERTATION DEFENSE

Fan Lai



Minimalist Systems for
Pervasive Machine Learning

Friday, June 23, 2023

1:30 – 3:30pm

3941 Beyster

Hybrid – [Zoom](#)

ABSTRACT: Although cloud computing has successfully fostered the last leap forward in machine learning (ML), today's ML is becoming increasingly unsustainable. First, the exponential growth in ML resource demand is outpacing the affordable growth in total resource capacity. Even worse, the conventional wisdom of collecting everything into the cloud and then improving ML is becoming infeasible, due to the skyrocketing volumes of edge data and tightening data restrictions (e.g., regulations, user privacy concerns).

This talk demonstrates how we can build a software systems stack that embraces minimalism at its core to overcome these two roadblocks. By co-designing ML, systems, and networking, we can (1) minimize the resource demand of ML by slashing the total amount of system execution needed to achieve the same ML accuracy; and (2) minimize data collection by effectively offloading ML to the planet-scale data source. Finally, I will outline my vision for making both ML and systems highly accessible, efficient, and automated for the upcoming decade.

CHAIR: Prof. Mosharaf Chowdhury