COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF MICHIGAN

DISSERTATION DEFENSE

Elisa Tsai

Machine Learning for Security and Beyond: From Threat Detection to Coreset Selection for Efficient Learning

Tuesday, July 22, 2025
9:00am – 11:00am
3725 Beyster / Hybrid – Zoom

ABSTRACT: The past decade has witnessed a Cambrian explosion in artificial intelligence, driven by breakthroughs in model architectures, advances in training paradigms and optimization techniques, growing computational power, and the availability of massive datasets. These forces have fueled a proliferation of downstream applications, from autonomous vehicles and virtual assistants to critical domains such as healthcare, finance, and cybersecurity. As machine learning systems are increasingly integrated into high-stakes, real-world settings, they face two fundamental and interrelated challenges: the scale and noisiness of raw data, and the inefficiency of learning pipelines that depend heavily on costly human annotation and resource-intensive training.

This dissertation addresses these challenges through two complementary directions: (1) developing machine learning systems for real-world security threat detection, and (2) designing data-efficient learning algorithms that minimize annotation and training costs, with applications in security and beyond.

In Part I of the dissertation, we develop ML-based systems to detect a wide range of Internet threats, including social engineering scams, dark patterns in online terms and conditions, network intrusion, and network interference. These systems integrate models such as large language models for text understanding, neural networks for classification, decision trees for interpretability, and clustering algorithms for unsupervised pattern discovery. They are grounded in large-scale empirical measurements across globally distributed platforms. In this talk, we highlight one such system, TermLens, which uses LLMs to surface unfair or unfavorable financial terms and conditions on shopping websites.

In Part II of the dissertation, we focus on data-efficient learning through coreset selection: the process of selecting a small and representative subset of training data that preserves model performance. We begin by identifying a challenge in domains like cybersecurity and medical imaging, where data difficulty often correlates with class membership. To address this, we introduce a separability metric and propose class-aware sampling strategies that improve robustness. We also present a label-free coreset method that uses deep clustering for pseudo-label generation and a pruning strategy to estimate data difficulty without relying on ground-truth labels, enabling scalable learning when labeled data is limited and annotation is costly.

CHAIR: Prof. Atul Prakash