# Numerical Analysis and Scientific Computing Seminar

## *Attacking neural networks with poison frogs: a theoretical look at adversarial examples in machine learning*

Thomas Goldstein
University of Maryland

**Abstract:** Neural networks solve complex computer vision problems with human-like accuracy. However, it has recently been observed that neural nets are easily fooled and manipulated by "adversarial examples," in which an attacker manipulates the network by making tiny changes to its inputs. In this talk, I give a high-level overview of adversarial examples, and then discuss a newer type of attack called "data poisoning," in which a network is manipulated at train time rather than test time. Then, I explore adversarial examples from a theoretical viewpoint and try to answer a fundamental question: "Are adversarial examples inevitable?"
Bio: Tom is an Assistant Professor at University of Maryland. His research lies at the intersection of optimization and distributed computing, and targets applications in machine learning and image processing. He designs optimization methods for a wide range of platforms. This includes powerful cluster/cloud computing environments for machine learning and computer vision, in addition to resource limited integrated circuits and FPGAs for real-time signal processing. Before joining the faculty at Maryland, he completed his PhD in Mathematics at UCLA, and was a research scientist at Rice University and Stanford University. He has been the recipient of several awards, including SIAMs DiPrima Prize, a DARPA Young Faculty Award, and a Sloan Fellowship.

Friday, March 22, 2019, 2:00 pm
Mathematics and Science Center: W301

Mathematics
Emory University