



THE UNIVERSITY OF
CHICAGO

DEPARTMENT OF STATISTICS

Statistics Colloquium

Data Science Institute/Computer Science/Statistics Joint Seminar

MAHDI SOLTANOLKOTABI

Associate Professor
University of Southern California

Towards Stronger Foundations for AI and its Applications to the Sciences

MONDAY, February 28, 2022, at 4:30 PM

In-Person: John Crerar Library, Room 390, and

Remote: [Live Stream](#) or Zoom (details sent via email announcement)

Host: Rebecca Willett

ABSTRACT

Despite wide empirical success, many of the most commonly used learning approaches lack a clear mathematical foundation and often rely on poorly understood heuristics. Even when theoretical guarantees do exist they are often too crude and/or pessimistic to explain their success in practical regimes of operation or serve as a guiding principle for practitioners. Furthermore, in many scenarios such as those arising in scientific applications they require significant resources (compute, data, etc.) to work reliably.

The first part of the talk takes a step towards building a stronger theoretical foundation for such nonconvex learning. In particular, I will focus on demystifying the generalization and feature learning capability of modern overparameterized learning where the parameters of the learning model (e.g. neural network) exceed the size of the training data. Our result is based on an intriguing spectral bias phenomena for gradient descent, that puts the iterations on a particular trajectory towards solutions that are not only globally optimal but also generalize well. Notably this analysis overcomes a major theoretical bottleneck in the existing literature and goes beyond the “lazy” training regime which requires unrealistic hyperparameter choices (e.g. very small step sizes, large initialization or wide models). In the second part of the talk I will discuss the challenges and opportunities of using AI for scientific applications and medical image reconstruction in particular. I will discuss our work on designing new architectures that lead to state of the art performance and report on techniques to significantly reduce the required data for training.