

CHAI Seminar Series

DATE: **Monday, February 28, 2022, 12:00-1:00 PM**

SPEAKER: **Matus Telgarsky, Ph.D.**
Assistant Professor, University of Illinois, Urbana-Champaign

TITLE: **Approximation, Optimization, and Generalization in Deep Networks**

REGISTRATION LINK:
<https://www.rit.edu/events/chai-research-talk-matus-telgarsky>



ABSTRACT: This talk will survey three areas of deep learning theory, ranging from classical to modern results. The first question is: what functions can deep networks approximate? Classically, it was shown they can approximate anything, but this gives no hints about efficient approximations and architecture choices, so I will also present more recent results about benefits of depth and other architectural considerations. Secondly, comes the question of generalization: why do deep networks achieve such good test error, despite achieving perfect fitting, or, arguably, overfitting? Here too I will survey classical results using generalization theory and point out their weaknesses, and point to more recent ideas, most of which are tied to optimization. The last topic is in fact optimization: since deep network training is non-convex, why do gradient methods work well? Here I will mainly discuss recent works, both discussing the near-initialization regime (the "neural tangent kernel"), and various preliminary works away from initialization, in the "feature learning regime".

BIO: Matus Telgarsky is an assistant professor at the University of Illinois, Urbana-Champaign. He studies machine learning theory, with a focus on deep learning theory and nonconvex optimization. He is an NSF CAREER award winner. He has significantly advanced the theory of deep learning, including the first proof demonstrating the benefits of depth in deep learning.