



# Vaishnavh Nagarajan

Explaining generalization in deep learning: progress and fundamental limits

Friday, July 30, 2021 – 3:00 p.m. – REMOTE

This dissertation studies a fundamental open challenge in deep learning theory: why do deep networks generalize well even while being overparameterized, unregularized and fitting the training data to zero error? In the first part of the thesis, we will empirically study how training deep networks via stochastic gradient descent implicitly controls the networks' capacity. Subsequently, to show how this leads to better generalization, we will derive *data-dependent uniform-convergence-based* generalization bounds with improved dependencies on the parameter count.

Uniform convergence has in fact been the most widely used tool in deep learning literature, thanks to its simplicity and generality. Given its popularity, we will also take a step back to identify the fundamental limits of uniform convergence as a tool to explain generalization. In particular, in the second part of the thesis, we will show that in some example overparameterized settings,  $\{\text{any}\}$  uniform convergence bound will provide only a vacuous generalization bound.

With this realization in mind, in the last part of the thesis, we will change course and introduce an empirical technique to estimate the generalization gap using unlabeled data. Our technique does not rely on any notion of uniform-convergence-based complexity and is remarkably precise. We will also theoretically show why our technique enjoys such precision.

We will conclude by discussing how future work could explore novel ways to incorporate distributional assumptions in generalization bounds (such as in the form of unlabeled data) and could explore other tools to derive bounds, perhaps by modifying uniform convergence or by developing completely new tools altogether.

**Thesis Committee:**

**J. Zico Kolter, Chair**

**Andrej Risteski**

**Ameet Talwalkar**

**Nathan Srebro Toyota Technological Institute at Chicago**