Stochastic memoization, Bayesian nonparametrics, and the goal of developing general-purpose predictive models.

Trying to design and learn general-purpose predictive models might seem crazy at first: for starters one might ask “What’s the point?” Also, one might ask, “How can one possibly hope to learn a good model without the benefit of deep, problem-specific insights?” In this talk I will argue that building general-purpose predictive models might be within the realm of possibility. I’ll start by reviewing general-purpose lossless compression (an application that benefits from fully-unsupervised, general-purpose predictive models), connect this to probabilistic modeling, then show how my group’s work has led to a new and powerful class of general purpose lossless compressors built on Bayesian nonparametric models; inference in which can be described as stochastic memoization. I will put my group’s work in the context of similar research being carried out at many different levels of generative model complexity and talk about future directions for both my own work and the field as a whole.