Dr. Gergely Neu

(Universitat Pompeu Fabra, Spain)

Title:

"Information-Theoretic Generalization Bounds for Stochastic Gradient Descent"

Abstract:

We study the generalization properties of the popular stochastic gradient descent method for optimizing general non-convex loss functions. Our main contribution is providing upper bounds on the generalization error that depend on local statistics of the stochastic gradients evaluated along the path of iterates calculated by SGD. The key factors our bounds depend on are the variance of the gradients (with respect to the data distribution) and the local smoothness of the objective function along the SGD path, and the sensitivity of the loss function to perturbations to the final output. Our key technical tool is combining the information-theoretic generalization bounds previously used for analyzing randomized variants of SGD with a perturbation analysis of the iterates.