



THE UNIVERSITY OF
CHICAGO

Department of Statistics

STATISTICS COLLOQUIUM

DANIEL HSU

Computer Science Department & Data Science Institute
Columbia University

Predictive Models from Interpolation

MONDAY, November 18, 2019 at 4:30 PM

Jones 303, 5747 S. Ellis Avenue

Refreshments before the seminar at 4:00 PM in Jones 304

ABSTRACT

Recent experiments with non-linear machine learning methods demonstrate the generalization ability of classification and regression models that interpolate noisy training data. It is difficult for existing generalization theory to explain these observations. On the other hand, there are classical examples of interpolating methods with non-trivial risk guarantees, the nearest neighbor rule being one of the most well-known and best-understood. I'll describe a few other such interpolating methods (old and new) with stronger risk guarantees compared to nearest neighbor in high dimensions. In some cases, we can demonstrate optimal rates of convergence with interpolating methods. This counters the conventional wisdom that interpolation necessarily implies poor generalization, and sheds light on the inductive biases that we rely on in machine learning.

For further information and inquiries about building access for persons with disabilities, please contact Jonathan Rodriguez at 773.702.8333 or send him an email at jgrodriquez@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following website:
<https://lists.uchicago.edu/web/subscribe/statseminars>.