

Stanford University
Department of Statistics

DEPARTMENTAL SEMINAR

4:30pm, Tuesday, October 16, 2018
Sloan Mathematics Center Room 380C

Refreshments served at 4pm in Sequoia Lounge.

Speaker: Tengyu Ma, *Stanford Statistics and Computer Science*

Title: **On the Margin Theory of Feedforward Neural Networks**

Abstract:

Past works have shown that, somewhat surprisingly, over-parametrization can help generalization in neural networks. Towards explaining this phenomenon, we adopt a margin-based perspective. We establish: 1) for multi-layer feedforward relu networks, the global minimizer of a weakly-regularized cross-entropy loss has the maximum normalized margin among all networks; 2) as a result, increasing the over-parametrization improves the normalized margin and generalization error bounds for two-layer networks. In particular, an infinite-size neural network enjoys the best generalization guarantees. The typical infinite feature methods are kernel methods; we compare the neural net margin with that of kernel methods and construct natural instances where kernel methods have much weaker generalization guarantees. We validate this gap between the two approaches empirically. Finally, this infinite-neuron viewpoint is also fruitful for analyzing optimization. We show that a perturbed gradient flow on infinite-size networks finds a global optimizer in polynomial time.