THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY

Department of Mathematics

SEMINAR ON DATA SCIENCE AND MACHINE LEARNING

How Much Over-parameterization Is Sufficient to Learn Deep ReLU Networks?

By

Dr. Yuan CAO
University of California at Los Angeles

Abstract

A recent breakthrough in deep learning theory shows that the training of over-parameterized deep neural networks (DNNs) can be characterized by the neural tangent kernel (NTK). However, existing optimization and generalization guarantees for deep neural networks typically require a network width larger than a high degree polynomial of the training sample size n and the inverse of the target accuracy ϵ^{-1}. In this talk, I will discuss how this over-parameterization condition can be improved to more practical settings. Specifically, I will first explain why over-parameterized DNNs can be optimized to zero training error in the NTK regime, and then show what kind of functions can be learnt by DNNs with small test errors. Under standard NTK-type assumptions, these optimization and generalization guarantees hold with network width polylogarithmic in n and ϵ^{-1}.

Biography: *Yuan Cao is a postdoctoral researcher in the Department of Computer Science at UCLA working with Professor Quanquan Gu. Before joining UCLA, he received his B.S. from Fudan University and Ph.D. from Princeton University. Yuan's research interests include the theory of deep learning, non-convex optimization, high-dimensional graphical models and their applications in computational genomics.*

Date : **30 April, 2020 (Thursday)**
Time : **3:00pm – 4:00pm**
Zoom Meeting : **https://hkust.zoom.com.cn/j/5616960008**

All are Welcome!