THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY

Department of Mathematics

SEMINAR ON DATA SCIENCE AND APPLIED MATH

A Generalized Neural Tangent Kernel Analysis for Two-layer Neural Networks

By

Dr. Yuan CAO
UCLA

Abstract

A recent line of research on deep learning shows that the training of extremely wide neural networks can be characterized by a kernel function called neural tangent kernel (NTK). However, it is known that this type of result does not perfectly match the practice, as NTK-based analysis requires the network weights to stay very close to their initialization throughout training, and cannot handle regularizers or gradient noises. In this talk, I will present a generalized neural tangent kernel analysis and show that noisy gradient descent with weight decay can still exhibit a ``kernel-like'' behavior. This implies that the training loss converges linearly up to a certain accuracy. I will also discuss the generalization error of an infinitely wide two-layer neural network trained by noisy gradient descent with weight decay.

Biography: *Dr. 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	:	**14 August 2020 (Friday)**
Time	:	**11:00am – 12:00noon**
Zoom Meeting	:	**https://hkust.zoom.us/j/5616960008**

All are Welcome!