Optimization Algorithms for Robust Machine Learning with Big Data

Yan Yan, University of Iowa

Date: Monday, March 2, 2020
Time: 9:10am-10:30am
Location: ETRL 101

Abstract:
Machine learning has been extensively used in many application areas, where big data imposes a variety of challenges. How can we learn a predictive model that is more robust with faster convergence speed? In this presentation, I would like to share some of my research works on developing accelerated optimization algorithms for enhancing robustness with theoretical analysis. I would start with an online learning algorithm for imbalanced data that is built on multiple cost-sensitive learners and makes use of "learning with expert advice" technique for online prediction. To justify the significance, a performance guarantee on F-measure of the proposed algorithm has been provided. Then I would present my effort towards faster convergence in terms of testing error for training deep neural networks by stochastic gradient descent (SGD), in which I provide both theoretical analysis and empirical evidence to verify the efficacy of stagewise SGD than the standard SGD. Finally, I would introduce my investigation on min-max optimization, where I develop a stagewise primal-dual algorithm to achieve faster convergence than $O(1^{\sqrt{T}})$ without assuming the bilinear structure over the objective function.

Bio:
Yan Yan received his Ph.D. in computer science from University of Technology Sydney, Australia in 2018 and his B.E. in computer science from Tianjin University, China in 2013. He is now a postdoctoral research associate at University of Iowa, USA. His research interests widely include theoretical methodology of statistical machine learning and its outreach on applications of computer vision and data mining. His current focus is on large-scale robust machine (deep) learning. In particular, he works on designing optimization algorithms to enhance the generalization performance for machine learning problems with various mathematical formulations, including minimization, min-max and inf-projection structures. In addition, his research projects also covered topics of online learning for imbalanced data, matrix factorization for recommender systems, etc.