NYU Stern School of Business
Department of Information, Operations & Management Sciences
STATISTICS RESEARCH SEMINAR

TOPIC: Analyzing Big Corrupted Data
SPEAKER: Xu Huan (National University of Singapore)
DATE: Friday, November 6, 2015
TIME: 11:30 AM - 12:30 PM
PLACE: KMC 4-80

Abstract

The explosive growth of data scale has presented great challenges to machine learning, demanding methods to handle large-scale, high-dimensional and dynamically changing data. Moreover, noise, data contamination and outliers, which are ubiquitous in realistic data, make robustness a necessary property for learning algorithms to succeed in practice. In this talk, I will discuss how to design robust learning algorithm that handles large scale data. This can be achieved by scaling up robust algorithms, which I illustrate with two concrete examples: Nystrom acceleration for robust PCA, and divide-and-conquer for graph clustering. This can also be achieved by robustifying large scale learning algorithms, as an example I will explain a general robust distributed learning framework based on geometric median.

Bio

Huan Xu received the B.Eng. degree in automation from Shanghai Jiaotong University, Shanghai, China, the M.Eng. degree in electrical engineering from the National University of Singapore, and the Ph.D. degree in electrical engineering from McGill University, Canada. From 2009 to 2010, he was a postdoctoral associate at The University of Texas at Austin. Since 2011, he has been an assistant professor at the Department of Mechanical Engineering at the National University of Singapore. His research interests include statistics, machine learning, robust optimization, and planning and control. He has published in premium venues of operations research and of machine learning, including Operations Research, Mathematics of Operations Research, Mathematical Programming, Journal of Machine Learning Research, IEEE Transaction of Information Theory, ICML, NIPS etc. He is an associate editor of IEEE Transactions on Pattern Analysis and Machine Intelligence and is on the editorial board of Computational Management Science.