NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

CS SEMINAR

Title: **Towards Practical Machine Learning with Differential Privacy and Beyond**

Speaker: Dr Yu-Xiang Wang
Scientist at Amazon AI

Date/Time: 15 August 2017, Tuesday, 01:00 PM to 02:00 PM

Venue: MR3, COM2-02-26

Chaired by: Dr Wang Wei, Assistant Professor, School of Computing
(wangwei@comp.nus.edu.sg)

Abstract:
Machine learning (ML) has become one of the most powerful classes of tools for artificial intelligence, personalized web services and data science problems across fields. However, the use of ML on sensitive data sets involving medical, financial and behavioral data are greatly limited due to privacy concern. In this talk, we consider the problem of statistical learning with privacy constraints. Under Vapnik's general learning setting and the formalism of differential privacy (DP), we establish simple conditions that characterizes the private learnability, which reveals a mixture of positive and negative insight. We then identify generic Bayesian learning methods that reuse existing randomness to effectively solve private learning in practice; and discuss a weaker notion of privacy ??? on-avg KL-privacy ??? that allows for orders-of-magnitude more favorable privacy-utility tradeoff, while preserving key properties of differential privacy. Moreover, we show that On-Average KL-P rivacy is **equivalent** to generalization for a large class of commonly-used tools in statistics and machine learning that sample from Gibbs distributions---a class of distributions that arises naturally from the maximum entropy principle. Finally, I will describe a few exciting future directions that use statistics/machine learning tools to advance he state-of-the-art for privacy, and use privacy (and privacy inspired techniques) to formally address the problem of p-hacking (or selective bias) in scientific discovery.

Biodata:
Yu-Xiang Wang is a scientist at Amazon AI. His research interests include statistical machine learning, deep learning, optimization and data privacy. Prior to joining Amazon, Yu-Xiang obtained his BEng and MEng in NUS, and spent four amazing years completing his PhD with the Machine Learning Department of Carnegie Mellon University. As an academic researcher, Yu-Xiang authored/co-authored numerous scientific articles on topics ranging from fundamental statistical methodology to internet-scale optimization and learning systems and produced applications in computer vision, recommender systems, web security, urban taxi analytics and even political campaigns. He received paper awards from KDD and WSDM, outstanding reviewer award at NIPS and the Baidu Scholarship among other

honors. Professionally, Yu-Xiang served on program committee for leading machine learning and AI conferences (ICML, NIPS, AISTATS) and reviewed dozens of papers for major journals in both statistics and computer sciences (JMLR, PAMI, Annals of Statistics).