

CS Seminar

Building Systems for Machine Learning

Dr Hong Xu
Department of Computer Science
City University of Hong Kong

Date:

August 8, 2019
Thursday
2:00 pm

Venue:

Room 313
Chow Yei Ching Building
The University of Hong Kong

Abstract:

Systems research is critical for machine learning because the recent success of AI and big data is in large part enabled by datacenter-scale computing infrastructures, which employ an army of machines to harness massive datasets in a continuous fashion. In this talk, I will present my research that focuses on systems for machine learning. First, we build a new distributed training system called Stanza that improves the training throughput of parameter server systems by 1.25x to 10.12x. Second, we build a serving system called Saec for recommendation models that reduces the memory footprint of embedding based recommendation models by 27x without performance loss.

About the Speaker:

Hong Xu is an associate professor in Department of Computer Science, City University of Hong Kong. His research area is computer networking and systems, particularly data center networks and big data systems. He received the B.Eng. degree from The Chinese University of Hong Kong in 2007, and the M.A.Sc. and Ph.D. degrees from University of Toronto in 2009 and 2013, respectively. He was the recipient of an Early Career Scheme Grant from the Hong Kong Research Grants Council in 2014. He received several best paper awards, including the IEEE ICNP 2015 best paper award. He is a senior member of IEEE and member of ACM.

All are welcome!
For enquiries, please call 2859 2180 or email
enquiry@cs.hku.hk
Department of Computer Science
The University of Hong Kong

