Yongjoo Park

http://yongjoopark.com
pyongjoo@umich.edu

INTERESTS

My research interests lie in **big data processing** and its applications to **data mining**. In particular, I focus on building *interactive* big data processing systems by leveraging machine learning and statistical techniques. This initiative has been applied to various applications, such as real-time data analytics, data visualizations, searching in high-dimensional space, etc., for which real-time answers could bring significant benefits to data analysts and decision makers.

EDUCATION

University of Michigan, Ann Arbor 2016
Ph.D. Candidate, Computer Science and Engineering
Thesis Advisors: Michael Cafarella and Barzan Mozafari

University of Michigan, Ann Arbor 2013
MS in Computer Science and Engineering
CGPA: 3.955/4.0

Seoul National University, Korea 2009
BS in Electrical Engineering

AWARDS

Graduate study (for PhD) scholarship 2013
Kwanjeong Educational Foundation
- The largest scholarship foundation in Korea

Graduate study (for Masters) scholarship 2011
Jeongsong Cultural Foundation
- One of eight students awarded in the year

National Science Scholarship 2004
Korea Student Aid Foundation (funded by Korea government)
- A full tuition support

PUBLICATIONS, PRESENTATIONS

Yongjoo Park, Amhad Shahab Tajik, Michael Cafarella, Barzan Mozafari
Database Learning: Toward a Database System that Becomes Smarter Over Time
In submission to The Proceedings of the Very Large Data Bases (PVLDB) 2016
Note: All three reviewers of this paper noted that "The paper will start a new line of research and products."

Yongjoo Park, Michael Cafarella, Barzan Mozafari
Visualization-Aware Sampling for Very Large Databases
IEEE International Conference on Data Engineering (ICDE) 2016
Note: A novel sampling method for one of the most frequently used data visualization method: scatter plot.

Yongjoo Park, Amhad Shahab Tajik, Michael Cafarella, Barzan Mozafari
Database Learning: Toward a Database System that Becomes Smarter Over Time
North East Database Day (NEDB) 2016, Oral, MIT
Note: A preliminary presentation for our VLDB paper above.

Yongjoo Park, Michael Cafarella, Barzan Mozafari
Neighbor-Sensitive Hashing
3rd Workshop on Web-scale Vision and Social Media (VSM) at ICCV 2015
Extended Abstract

Yongjoo Park, Michael Cafarella, Barzan Mozafari
Neighbor-Sensitive Hashing
The Proceedings of the Very Large Data Bases (PVLDB) 2015
Note: A significant improvement on an extremely well-known problem over the numerous works of the past decade.

Michael Anderson, Dolan Antenucci, Victor Bittorf, Matthew Burgess, Michael Cafarella, Arun Kumar, Feng Niu, **Yongjoo Park**, Christopher Ré, Ce Zhang
(authors in alphabetic order)
Brainwash: A Data System for Feature Engineering
The biennial Conference on Innovative Data Systems Research (CIDR) 2013

TEACHING EXPERIENCE

EECS 485 Web Databases and Information Systems ‘12 Winter
- Graduate Student Instructor, University of Michigan, Ann Arbor
- Designed programming assignments (interactive web using JavaScript, and PageRank computation of Wikipedia pages using Hadoop)
- Taught 100 students in weekly discussion sections

WORK EXPERIENCE

Software Engineer Internship, Amazon.com, Seattle ‘14 Summer
- Working in a Web team, I developed a data center capacity prediction system.

Software Engineer (Full-time), Webcash, Seoul 2009 - 2011
- Internet-banking project with J.P. Morgan Hongkong
- Financial iPhone application developments

Research Assistant, System Electronics Lab 2007
Seoul National University, Seoul
- Developed a power-efficient vehicle entertainment system that runs on embedded-processors (ARM)

SERVICE

External reviewers for VLDBJ’16, VLDB’16, VLDB’15, SIGMOD’16, ICDE’15, CIDR’15

Organizers of University of Michigan DB Group meetings (’16, ’14) and MIDAS (Michigan Data Science) seminars (’14)