Jason (Hyun Duk) Cho

CONTACT INFORMATION

University of Illinois at Urbana-Champaign
Computer Science Dept
Urbana, IL 61801, USA

Phone: +1-217-898-6642
E-mail: hcho33@illinois.edu
WWW: http://thejason.co

RESEARCH INTERESTS

Machine Learning, Information Retrieval, Natural Language Processing, Data Mining, Health Informatics

EDUCATION

University of Illinois at Urbana-Champaign, Urbana, IL

- Ph.D. Candidate in Computer Science — Since Fall 2012
  Advisor: Professor Roxana Girju, Professor Bruce R. Schatz

- M.S. in Computer Science — Fall 2011 – Summer 2012
  Thesis: *Utilizing multiple entities from collection of unstructured documents in constructing attribute-value pairs*
  Advisor: Professor Chengxiang Zhai

- B.S. in Computer Science — August 2007 – May 2011
  Graduated with Highest Honors (University Honors, GPA: 3.95)

INDUSTRY EXPERIENCE

- **Data Scientist Intern**, @WalmartLabs — May – August 2014
  - Supervisor: Shyam Rapaka
  - Researched and implemented personalized algorithms to improve walmart.com front page item recommendation module. Algorithm performed 10% better than the baseline on offline evaluations.
  - Designed the front page personalized item recommendation module architecture.
  - Currently coordinating both QA and AB testing on walmart.com.
  - Working on two patent applications.
- **Software Development Engineer Intern**, Amazon.com — May – August 2012
  - Supervisor: Samir Khobragade, Mentor: Suketu Vakharia
  - Designed and integrated comparison page of Amazon wireless.
  - Researched and developed carrier plan filter engine that would allow future machine learning techniques to be used.
- **Software Engineer Intern**, Jump Trading, LLC — June – August 2011
  - Supervisor: Stephen Yi
  - Designed and developed price feed system, which consisted of two parts, client library and price feed handler.
  - Integrated price feed and gateway system into Click-Trade UI.

RESEARCH EXPERIENCE

- **Research Assistantship**, University of Illinois at Urbana-Champaign 01/2013 – 05/2013,
  - Advisor : Prof. Schatz
  - Generate medically valid hypotheses (Comparative Effectiveness Research) by utilizing medical web forums.
  - Online medical forum post recommender systems (Case retrieval system).
- **Research Programmer**, Center for Simulation of Advanced Rockets 03/2009 – 05/2011
  - Advisor : Dr. Shaffer
  - Researched in classifying LiDAR data point by utilizing Markov Random Fields.
  - Implemented GPGPU sorting algorithm.

PUBLICATIONS

- Thomas Zhang, **Jason H.D. Cho**, Chengxiang Zhai: *"Understanding user Intents in Online Health Forums"*, In Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics (ACM BCB'14)
- **Jason H.D. Cho**, Parikshit Sondhi, Chengxiang Zhai, Bruce R. Schatz: *"Resolving Healthcare Forum Posts via Similar Thread Retrieval"*, In Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics (ACM BCB'14)
- **Jason H.D. Cho**, Jacob Guggenheim, Elizabeth Arcan, Hannah Friedman, Shravan Gupta, Joyce Thomas, Deana McDonagh, Bruce R. Schatz, *"Utilizing Smartphones to Enhance Urine Strip Accessibility" (Poster Paper)*, In Proceedings of American Medical Informatics Association Annual Symposium (AMIA '14)
- Manish Gupta, Arun Mallya, Subhro Roy, **Jason H.D. Cho**, Jiawei Han: *"Local Learning for Mining Outlier Subgraphs from Network Datasets"*, In Proceedings of 2014 SIAM Data Mining Conference (SDM '14)
- **Jason H.D. Cho**, Vera Q.Z. Liao, Yunliang Jiang, Bruce R. Schatz: *"Aggregating Personal Health Messages for Scalable Comparative Effectiveness Research"*, In Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics (ACM BCB'13)
- Mahsa Kamali, Matei Stroila, **Jason Cho**, Eric Shaffer, and John C. Hart : *"Robust Classification of Curvilinear and Surface-like Structures in 3d Point Cloud Data"*, In Proceedings of ISVC 2011, Journal of Lecture Notes on Computer Science, Springer Verlag

Teaching Experience

- **Teaching Assistant**, CS 225 (Data Structures) — June – August 2013
- **Teaching Assistant**, CS 440 (Artificial Intelligence) — August – December 2012
- **Course Grader**, CS 440 (Artificial Intelligence) — January – May 2011
- **Undergrad. Teaching Assistant**, CS 225 (Data Structures) — June – August {2009, 2010}
- **Undergrad. Teaching Assistant**, CS 125 (Intro. to CS) — August 2008 – December 2009

Mentoring Experience

- Topic: **Gamifying health data collection process** — February 2013 – Current
  Mentees: Mariko Wakabayashi (BS), Seungchul Lee (BS), Ronit Chakraborty (BS), Dong Min Shin (BS), RJ Kunde (BS), Saubhagya Rathi (BS), Thomas Olson (BS), Rafael Drummond (BS)
  Top ten finalists for CIMIT, won $\$10,000$ *(Spring 2014)*
  Won Promoting Undergraduate Research in Engineering (P.U.R.E) Audience Award (Spring 2013)
- Topic: **Classification of patient's intents on medical web forums** — August 2013 – May 2014
  Mentee: Thomas Zhang (5 year BS/MS Student)
- Topic: **Utilizing forum meta-information to improve relevance in forum discovery**
  Mentee: Han-wen Yeh (5 year BS/MS Student) — September 2012 – April 2013
- Topic: **Bias Detection in Journalism** — February 2013 – May 2013
  Mentees: Richard Lee (BS), James Chen (BS)

Invited Talks

- **Addressing Users Healthcare Needs through Personal Health Messages**, Data and Information Systems (DAIS) Seminar, UIUC — October 2013

Academic Activities

- **Book Reviewer**
  Learning Data Mining with R, Packt Publishing
- **Journal and Conference Reviewer**
  American Medical Informatics Association (AMIA) 2015
  World Wide Web (WWW) 2015
  Language and Linguistics Compass (LNCO) 2014

Honors and Awards

- Top ten finalists for Center for Integration of Medicine and Innovative Technology (CIMIT).
  Served as research mentor and project manager. Won $\$10,000$. — Spring 2014
- Promoting Undergraduate Research in Engineering (PURE) Awards
  Served as graduate student mentor. Mentees won audience award (Spring 2013) and best poster award

(Spring 2014)

- Health Information Technology Center (HITC) Fellow — Spring 2013 – Fall 2013
- HITC travel grant for American Medical Informatics Association (AMIA) — Fall 2012
- Jump Trading Scholarship — Fall 2010 – Spring 2012
  $15,000 for Fall 2010 – Spring 2011, $30,000 and tuition waiver for Fall 2011 – Spring 2012
- Bronze Tablet — Spring 2011
  Given to students whose GPA rank top 3% in their graduating class
- James N. Snyder Outstanding Undergraduate Award — Spring 2009
- Dean's List — Fall 2007 – Fall 2010

Relevant Courseworks

- CS 598 – Machine Learning for Signal Processing — Fall 2013
- CS 598 – Vision, Learning, AI & NLP — Spring 2013
- CS 512 – Data Mining Principles — Spring 2013
- MATH 484 – Nonlinear Programming — Fall 2012
- CS 598 – Health Informatics — Fall 2012
- CS 598 – Advanced NLP: Theory and applications of Bayesian models — Spring 2012
- CS 466 – Introduction to Bioinformatics — Spring 2012
- CS 598 – Advanced Topics in Information Retrieval — Fall 2011
- CS 598 – Probabilistic Graphical Model in AI — Fall 2011
- STAT 542 – Statistical Learning — Fall 2011
- MATH 482 – Linear Programming — Spring 2011
- CS 546 – Machine Learning in NLP — Spring 2011
- CS 410 – Text Information Systems — Spring 2011
- CS 446 – Machine Learning — Fall 2010
- CS 498 – Intro to Natural Language Processing — Fall 2010
- CS 412 – Intro to Data Mining — Fall 2010
- CS 440 – Intro to Artificial Intelligence — Spring 2010
- STAT 410 – Statistics and Probability II — Spring 2010
- STAT 420 – Methods of Applied Statistics — Spring 2010

Services and Memberships

- **CS Graduate Ambassador**, Computer Science, UIUC — Fall 2011 – May 2012
- **Alumni Advisor**, Triangle Fraternity, UIUC — Since January 2012
- **Student Member**, Association for Computing Machinery (ACM) — Since January 2013
- **Student Member**, ACM SigBio — Since July 2013
- **Student Member**, American Medical Informatics Association (AMIA) — Since October 2012

Skills

- **Programming Languages:** Java, C/C++, Matlab, R, Python, JavaScript, JQuery, HTML, CSS
- **OS/Environments:** Eclipse, Vim, Linux, Windows
- **Languages:** English (fluent), Korean (fluent), Chinese (intermediate)