

Songbai Yan

Email: yansongbai001@gmail.com

Cellphone: (858)766-1725

EDUCATION

Ph.D. Student in Computer Science, University of California, San Diego

Sep. 2014 - Present

Research Area: Machine Learning

B.S. in Computer Science, Peking University, Beijing, China

Sep. 2010 - Jun. 2014

Major GPA: 3.93, Rank: 4/150

EXPERIENCE

Active Learning with Abstention Feedback

University of California, San Diego

Graduate Student Researcher

Sept. 2014 - Present

- Proposed a novel active learning setting where the labeler can not only return noisy labels but also abstain from giving labels.
- Designed an algorithm with nearly optimal query complexity.

Cohort Selection with Presto

Hulu, Santa Monica

Software Developer Intern

Jun. 2015 - Aug. 2015

- Designed and implemented several user-defined functions in Presto for cohort selection based on user temporal behavior data.
- Built a user-friendly query interface using Bootstrap and AngularJS for the marketing team.
- Developed RESTful APIs with flask to process user input from the query interface.

Privacy-Preserving Data Analysis

Peking University

Undergraduate Student Researcher

Nov. 2013 - May 2014

- Improved accuracy of differentially private data releasing for smooth queries.
- Designed a privacy-preserving mechanism for low dimensional data description.

Learning Tractable Structures of Markov Networks

University of California, Los Angeles

Summer research intern. Funded by UCLA CSST Scholarship

Jul. 2013 - Sept. 2013

- Employed the Sentential Decision Diagram, a recently proposed representation scheme, for tractable bottom-up structure learning of Markov Networks.
- Presented the results in UCLA Cross-disciplinary Scholars in Science and Technology (CSST) program.

SELECTED COURSE PROJECTS

A Robust and Scalable Distributed Microblogging System

Spring, 2015

- Implemented scalable stateless front-ends and key-value store back-ends using RPC in Go.
- Replicated back-ends for fault-tolerance with distributed hash tables and two-phase commit.

PUBLICATIONS

S. Yan, K. Chaudhuri and T. Javidi. Active Learning from Imperfect Labelers. NIPS 2016

S. Yan, K. Chaudhuri and T. Javidi. Active Learning from Noisy and Abstention Feedback. Allerton 2016

K. Fan, H. Zhang, **S. Yan**, L. Wang, W. Zhang, J. Feng. Learning a Generative Classifier with Label Proportions. Neurocomputing, 2014

SELECTED HONORS AND AWARDS

UCSD Fall 2014 Programming Contest, 1st place

Oct. 2014

Peking University Tencent Innovative Scholarship (2 out of 150)

Sept. 2012

ACM-ICPC Asia Regional Contest Dalian Site, Silver Medal (10/100 in the top programming contest in Asia)

Sept. 2011

SKILLS

- Algorithm design, analysis and implementation
- Machine learning algorithms and theories
- Programming Languages: C/C++, Python, Java