

EDUCATION

- University of California, Berkeley, CA** *2019 – present*
Ph.D., Computer Science, Advised by Aditya Parameswaran
- University of Illinois Urbana-Champaign, IL** *2014 – 2019*
M.S., Ph.D. (Transferred), Computer Science, Advised by Aditya Parameswaran
- California Institute of Technology, Pasadena, CA** *2008 – 2012*
B.S., Computer Science, *Graduated with Honors.*
-

HONORS & AWARDS

- Heidelberg Laureate Forum Participant** *2019*
One of 200 young researchers world-wide invited to participate at the Heidelberg Laureate Forum.
- SIGMOD Student Travel Award** *2018*
Awarded by NSF to support students attending the ACM SIGMOD/PODS conference.
- NSF GRFP Fellowship** *2016*
Three-year stipend and full-tuition waiver awarded by the National Science Foundation to domestic students in STEM. Acceptance rate: 11.8%.
- DREAM Prostate Cancer Challenge Task 1A Winner** *2015*
Computational challenge on predicting the survival of patients with metastatic castrate resistant prostate cancer using genomics data.
- Conference Travel Award for Graduate Students** *2014*
Awarded by UIUC Graduate College to support students presenting at professional conferences.
- State Farm Companies Foundation Doctoral Scholar** *2014*
Awarded by UIUC CS Department to promising incoming doctoral students.
- Richard T. Cheng Fellowship** *2014*
Awarded by UIUC CS Department to exceptional incoming graduate students.
- TA Excellence Award** *2012*
Awarded by the Caltech CMS Department for being an effective TA based on student feedback.
- Grace Hopper Scholarship** *2011*
Awarded by Facebook to attend the Grace Hopper Celebration of Women in Computing 2011.
-

WORK/RESEARCH EXPERIENCE

- Graduate Research Student** *August 2019 – Present*
University of California, Berkeley, CA
- Research Intern** *May 2019 – August 2019*
Microsoft, Cloud and Information Systems Lab (CISL), Sunnyvale, CA
- Graduate Research Fellow** *August 2016 – May 2019*
University of Illinois Urbana-Champaign, IL
- .406 Ventures Student Fellow** *September 2018 – Present*
.406 Ventures, Boston, MA
- Research Intern** *May 2016 – August 2016*
Google Research, Mountain View, CA

Research Assistant <i>Data Mining Research Group, University of Illinois Urbana-Champaign, IL</i>	<i>August 2014 – May 2016</i>
Software Engineering Intern <i>Google Inc., Mountain View, CA</i>	<i>May 2015 – August 2015</i>
Software Engineering Intern <i>Databricks, Berkeley, CA</i>	<i>May 2014 – August 2014</i>
Senior Software Engineer <i>Recommendations Team, LinkedIn, Mountain View, CA</i>	<i>March 2014 – May 2014</i>
Software Engineer <i>Recommendations Team, LinkedIn, Mountain View, CA</i>	<i>July 2012 – March 2014</i>
Senior Thesis <i>Winfree Lab, California Institute of Technology, CA</i>	<i>September 2011 – June 2012</i>
Research Intern <i>SELECT Lab, Carnegie Mellon University, Pittsburgh, PA</i>	<i>June 2011 – September 2011</i>
Research Intern <i>Machine Learning and Instrument Autonomy Group, Jet Propulsion Lab</i>	<i>June 2010 – August 2010</i>
Undergraduate Research Fellow <i>Jensen Lab, California Institute of Technology, Pasadena, CA</i>	<i>June 2009 – September 2009</i>

TEACHING EXPERIENCE

Guest Lecturer <i>Database Systems, UIUC</i>	<i>October 2018</i> with Professor Abdu Alawini
Teaching Assistant <i>Coursera MOOC: Principles of Data Mining, UIUC</i>	<i>January 2015 – May 2015</i> with Professor Jiawei Han
Teaching Assistant <i>CS/EE/Ma 156a: Learning from Data, Caltech</i>	<i>January 2012 – March 2012</i> with Professor Yaser Abu-Mostafa
Teaching Assistant <i>CS/EE/Ma 129a: Information and Complexity, Caltech</i>	<i>September 2011 – December 2011</i> with Professor Erik Winfree

PUBLICATIONS

Extending Relational Query Processing with ML Inference. Konstantinos Karanasos, Matteo Interlandi, **Doris Xin**, Fotis Psallidas, Rathijit Sen, Kwanghyun Park, Ivan Popivanov, Supun Nakandal, Subru Krishnan, Markus Weimer, Yuan Yu, Raghuram Ramakrishnan, Carlo Curino. *The Conference on Innovative Data Systems Research (CIDR)*, January 2020.

A Human-in-the-loop Perspective on AutoML: Milestones and the Road Ahead. Doris Jung-Lin Lee*, Stephen Macke*, **Doris Xin***, Angela Lee, Silu Huang, Aditya Parameswaran (* equal contribution). *IEEE Data Engineering Bulletin, Issue on DB4AI and AI4DB*, June 2019

Helix: Holistic Optimization for Accelerating Iterative Machine Learning. **Doris Xin**, Stephen Macke, Litian Ma, Jialin Liu, Rong Ma, Aditya Parameswaran. *45th International Conference on Very Large Data Bases (VLDB)*, 2019

Active Learning on Heterogeneous Information Networks: A Multi-armed Bandit Approach. **Doris Xin**, Ahmed El-Kishky, De Liao, Brandon Norick, and Jiawei Han. *IEEE International Conference on Data Mining*, 2018

How Developers Iterate on Machine Learning Workflows — A Survey of the Applied Machine Learning Literature. **Doris Xin**, Litian Ma, Shuchen Song, Aditya Parameswaran. *KDD Workshop on Interactive Data Exploration and Analytics (IDEA)* (oral presentation), 2018

Helix: Accelerating Human-in-the-loop Machine Learning (Demo). **Doris Xin**, Litian Ma, Jialin Liu, Stephen Macke, Shuchen Song, Aditya Parameswaran. *44th International Conference on Very Large Data Bases (VLDB)*, 2018

Accelerating Human-in-the-loop Machine Learning: Challenges and Opportunities (Vision Paper). **Doris Xin**, Litian Ma, Jialin Liu, Stephen Macke, Shuchen Song, Aditya Parameswaran. *Proceedings of the 2nd Workshop on Data Management for End-to-End Machine Learning (DEEM)*, 2018

Folding: Why Good Models Sometimes Make Spurious Recommendations. **Doris Xin**, Nicolas Mayoraz, Hubert Pham, Karthik Lakshmanan, John Anderson. *Proceedings of the Eleventh ACM Conference on Recommender Systems (RecSys)*, 2017

DWCox: A density-weighted Cox model for outlier-robust prediction of prostate cancer survival. J. Xiao, S. Wang, J. Shang, H. Lin, **D. Xin**, X. Ren, J. Han, J. Peng. *F1000Research*, 2016

MLlib: Machine Learning in Apache Spark. X. Meng, J. Bradley, B. Yuvaz, E. Sparks, S. Venkataraman, D. Liu, J. Freeman, D. Tsai, M. Amde, S. Owen, **D. Xin**, R. Xin, M. Franklin, R. Zadeh, M. Zaharia, A. Talwalkar. *Journal of Machine Learning Research*, 2015

Parallel computation using active self-assembly. Moya Chen, **Doris Xin**, Damien Woods. *Natural Computing* 14, 2015

Active Learning in Heterogeneous Information Networks. **Doris Xin**, De Liao, and Jiawei Han. *Proceedings of International School and Conference on Network Science*, 2015.

LASER: A Scalable Response Prediction Platform for Online Advertising. Deepak Agarwal, Bo Long, Jonathan Traupman, **Doris Xin** and Liang Zhang (name listed in alphabetical order). *ACM WSDM*, 2014.

Metronome: Building Blocks for Data Products. Paul Ogilvie, Jonathan Traupman, Xiangrui Meng and **Doris Xin**. *BigData Innovators Gathering*, 2014.

Parallel Computing Using Active Self-assembly. Moya Chen, **Doris Xin** and Damien Woods. *DNA Computing and Molecular Programming*, 2013. *Best Student Paper Award*

INVITED TALKS

Accelerating Machine Learning Development through Automation.
Google. Mountain View, CA. 2020.

Towards Machine Learning as a Turnkey Technology.
Target AI Research, Sunnyvale, CA. 2019.

MLLib in Apache Spark. Insight Fellowship Program Invited Speaker. 2014.

SERVICE

Member of the UIUC Engineering Graduate Student Advisory Committee. 2018-2019

NSF GRFP Application Internal Reviewer for UIUC Applicants. 2017, 2018, 2019

Student Volunteer at ACM SIGMOD. 2017

Research Mentor to Undergraduate Students. 2015 - 2019

Graduate Ambassador to Prospective PhD Students. 2015 - 2019

Graduate Mentor to New PhD Students. 2015 - 2017

TECHNICAL SKILLS

Languages: Java, Scala, Python, Pig, C++, JavaScript, PHP, L^AT_EX

Frameworks/Software: Apache Hadoop, Apache Spark, FlumeJava, ScikitLearn, MySQL, Eclipse, IntelliJ, Git, SBT, Maven, Gradle, Matlab, Bootstrap, Tornado, d3
