

Xide Xia

Curriculum Vitae

MCS138, 111 Cummington Mall
Boston, MA 02215

Phone: (401) 209-4920
Email: xidexia@bu.edu

EDUCATION

Boston University

Ph.D. candidate in Computer Science

- *Dean's Fellow Scholarship 2016-2017*

2016 - Present

Harvard University, Institute for Applied Computational Science

M.E. in Computational Science and Engineering

- *Harvard IACS Student Scholarship 2015*

2014 - 2016

Brown University, School of Engineering

M.S. in Electrical Science

2012 - 2013

Beijing Institute of Technology, College of Information and Electronics

B.S. in Electrical and Information Engineering

- *Renmin Scholarship* (2008-2012, awarded to students with academic excellence and research potential)

2008 - 2012

PUBLICATION

2016

- Xide Xia, Mohammed AlQuraishi. "Computational prediction of protein-DNA interactions based on protein sequences information." Forthcoming.
- Xide Xia, Finale Doshi-Velez, Pavlos Protopapas. "Cost-Sensitive Batch Mode Active learning: Designing Astronomical Observation by Optimizing Telescope Time and Telescope Choice." In *In Proceedings of SIAM Data Mining Conference (SDM)*. 2016.

2015

- AlQuraishi, M., Tang, S., & Xia, X. (2015). An affinity-structure database of helix-turn-helix: DNA complexes with a universal coordinate system. *BMC bioinformatics*, 16(1), 390. PMID:26586237. (Database: <http://staging.proteindna.hms.harvard.edu/>).

RESEARCH EXPERIENCE

Boston University, Boston, MA

September 2016 – Present

Research Fellow in Image and Video Computing (IVC) Lab.

Harvard Medical School, Boston, MA

June 2013 – May 2016

Graduate Research Fellow advised by Dr. Mohammed AlQuraishi and Professor Peter Sorger.

- Develop a new computational method for predicting protein-DNA interactions by implementing a whole machine learning pipeline for prediction protein-DNA interactions.
- Build a model to simulate protein-DNA interactions, for which we formulate the problem as a convex optimization problem. In the model, we measure energies of binding between protein and DNA molecules to derive the energies of interactions between atoms.
- Impose all the constraints and implement a regression problem that is specific to the formulation.
- Predict protein-DNA interactions based on protein sequence even if the structure information is missing (i.e. CIS-BP database).
- Implement large scale scientific computing in parallel and distributed environments.
- Assist Dr. AlQuraishi to construct an affinity-structure database in which structural and biochemical data of protein-DNA interactions are integrated:
<http://staging.proteindna.hms.harvard.edu/>.

Harvard University, Cambridge, MA

Intervention and Outcome Predictions: Understanding Vasopressor Administration and Weaning in the ICU

Oct 2015 – Present

Research advised by Professor Finale Doshi-Velez.

- Analyze the MIMIC II 2.6 database gathered from four ICUs at the Beth Israel Deaconess Medical Center.
- Design a recurrent neural network model to simulate multidimensional physiological time series of patients before, during, and after vasopressor administration.
- Implement recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) units in Torch-Autograd.
- Predict whether a patient will be administered a vasopressor.

Batch Mode Active Learning and Its Application to Astronomy

Feb 2015 – Nov 2015

Thesis project advised by Professor Finale Doshi-Velez and Dr. Pavlos Protopapas.

- Developed a batch-mode cost-sensitive active learning approach that not only exploited uncertainty and representativeness of the whole unlabeled dataset, but also took annotation cost into consideration.
- Designed a selection criterion that combined uncertainty and representativeness by using a synthesized heuristic argument. Queried instances were selected from different clusters to have a good representativeness of the whole unlabeled dataset. The selection criterion also considers the overall cost of the experiment while keeping uncertainty and representativeness high.
- Applied the approach to optimize astronomical observations for object classification, including an extension that incorporates the possibility that nearby objects is observed at the same time.
- Used two large astronomical data sets (MACHO and EROS) to demonstrate how the approach balanced the trade-off among FOV, aperture, and time cost.

TEACHING EXPERIENCE

Boston University, Boston, MA

- CS131 Combinatoric Structures

2016 Fall

Harvard University, Cambridge, MA

- AM207 Stochastic Methods for Data Analysis, Inference and Optimization

2016 Spring

WORK EXPERIENCE

Boston University, Boston, MA

Research fellow at IVC Lab

Sept 2016 – Present

Harvard Medical School, Boston, MA

Research fellow at Department of Systems Biology

June 2013 – May 2016

Huawei Technologies Co., Ltd, Shenzhen, China

Software development engineer intern

June 2012- Aug 2012

Agilent Technologies Co., Ltd, Beijing, China

Undergraduate intern at mobile broadband division

Sept 2011- Jan 2012