Xi He

※ xih314@lehigh.edu ※ (484)633-8040 ※ 837 Cedar Hill Drive, Allentown, PA 18109 ※ www.xihey.com

To pursue a **research/quantitative/data scientist** position

Education

Ph.D.	Industrial and Systems Engineering, Lehigh University	Aug 2014 - Dec 2018 (Exp)
B.S. & M.S.	Mathematics (GPA ∼3.8/4.0), Nankai University, China	Sep 2008 - May 2014

Qualifications and Skills

Machine Learning	**C**lassification, **R**egression, **C**lustering, **N**eural networks (Dnn/Cnn/Rnn/Drl/Gan)
Optimization	**N**onlinear optimization and algorithms design, **E**xperience in integer programming
Industry	**Q**uantitative research in e-Trading equities, **M**arketing model
Programming	**D**istributed and parallel computing, **O**bject-oriented design, **V**isualization
	Python(Proficient), C++(Proficient), Matlab, R, Mathematica, Sql, Shell Script, Html
	Tensorflow/Pytorch, Mpi/OpenMP, Git, Aws, Gurobipy, Spark, Cuda, Linus, LaTeX

Working Experience

Quantitative Research Summer Associate *e-Trading Equities, LQR, JPMorgan Chase & Co.*, New York

- Designed deep reinforcement learning framework for optimal Algo execution problem. Python, Java, kdb+/q
- Achieved consistent reduction of aggressive ratio cross markets on deep Q-learning agent for given parent orders.
- Developed actor-critic agent and placement regulations for exploring uncertain continuous action space. **Jun-Sep, 2018**

Data Scientist Intern *Alliance Data, Precima Division, R&D Dept.*, Chicago

- Built appropriate large scale integer model for budgeted category offer assignment optimization. Python, Sql, Odbc
- Designed and implemented Lagrangian Relaxation approach with efficient greedy heuristic for model solving. Gurobipy
- Proposed distributed multi-threads algorithm to solve up-to 1M customers case in less than 1.5hrs. Mpi4py **May-Aug, 2017**

Data Scientist Intern *Siemens Corporation, Corporate Research*, Princeton

- Constructed multi-agent reinforcement learning model for Job-shop scheduling problems regarding make-span minimization.
- Designed appropriate observation and reward for solving the RL model with performance comparison. Python **Feb-May, 2017**
- Developed scalable subsample Hessian-free trust region framework for training deep feedforward neural networks. Python, Theano, Matlab
- Built ϵ-approx second-order convergence guarantee for proposed algorithms integrating negative curvature direction. **Jun-Dec, 2016**
- Proposed semi-positive definite quadratic programming model to measure default risk of a portfolio using normal copula. R
- Applied importance sampling to attain reliable and stable loss probabilities given optimal risk loading. **Jun-Aug, 2015**

Research Experience

Research Assistant *Lehigh University*

★ **Major Subject** Distributed Algorithms in Large-scaled Convex/nonconvex Empirical Risk Minimization. **Advisor: Martin Takáč**

- Proposed dual free adaptive mini-batch SDCA for empirical risk minimization for both convex and nonconvex loss. C++
- Developed large-scaled distributed Hessian-free methods for training deep fully-connected and convolutional neural network. C++
- Proved the best worst-complexity result of an purposed inexact regularized stochastic newton method for non-convex optimization.
- Performed distributed restarting Newton-CG method for large-scale empirical risk minimization. Python

Selected Courses and Projects

Massive Data Mining, Pattern Recognition, Natural Language Processing

- Designed question & answer system for 8th grade problems, attaining 39.5% accuracy using Apache Lucene etc. Python **Spring 2016**
- Released a Matlab package using various classifiers (SVM, KNN, etc.) for character-image recognition. Matlab **Fall 2015**

Computation Method, Optimization in Machine Learning, Nonlinear Programming

- Using ℓ_1-regularized lasso model to recovery pictures with missing pixels, applied ISTA, FISTA and GPSR algorithms. C++ **Fall 2015**
- Built a Matlab package for unconstrained nonlinear optimization applied typical algorithms and did comparison. Matlab **Spring 2015**

Publications

- **He Xi**, Rachael Tappenden, and Martin Takác. *Dual Free Adaptive Mini-batch SDCA for Empirical Risk Minimization.* Frontiers in Applied Mathematics and Statistics, 2017.
- **He Xi**, Dheevatsa Mudigere, Mikhail Smelyanskiy, and Martin Takáč. *Large Scale Distributed Hessian-Free Optimization for Deep Neural Network.* AAAI 2017 Workshop on Distributed Machine Learning.
- **He Xi** and Martin Takáč. *Dual Free SDCA for Empirical Risk Minimization with Adaptive Probabilities.* NIPS 2015 Workshop on Optimization in Machine Learning.
- Li YiYong, QingZhi Yang, and **He Xi**. *A Method with Parameter for Solving the Spectral Radius of Nonnegative Tensor.* Journal of the Operations Research Society of China, 2017.