Yining Wang

CONTACT INFORMATION

Room 8021, GHC Building
Carnegie Mellon University
Pittsburgh, PA, USA

Mobile: +1(412)721-3814
E-mail: ynwang.yining@gmail.com
WWW: yining-wang.com

EDUCATION

Carnegie Mellon University, Pittsburgh, PA, United States of America

PhD student, Machine Learning Department, since July 2014

- Graduate GPA: 4.05/4.3
- Courses: Advanced introduction to machine learning, intermediate statistics, statistical machine learning, convex optimization, advanced statistic theory, high-dimensional statistics, foundations of machine learning.
- Advisor: Prof. Aarti Singh
- Research interest: structured matrix learning, missing data, interactive learning, spectral methods, hashing.

Tsinghua University, Beijing, China

B.Eng., Computer Science and Technology, June 2014

- Enrolled in Yao Class, a pilot CS class directed by Prof. Andrew Chi-Chih Yao.
- GPA: 93.5/100; last two-year: 94.5/100; class ranking: 2/35
- Undergradaute thesis: *Spectral Methods in Supervised Topic Modeling* (in Chinese)
- Thesis advisor: Prof. Jun Zhu

Massachusetts Institute of Technology, Cambridge, MA, United States of America

Exchange student in Dept. of EECS, February 2013 to May 2013

- GPA: 5.0/5.0
- Courses: Inference and Information, Nonlinear Programming and Automatic Speech Recognition

PREPRINTS

[1] Yining Wang, Adams Wei Yu and Aarti Singh. **Computationally Feasible Near-Optimal Subset Selection for Linear Regression under Measurement Constraints**. *arXiv:1601.02068*.

[2] Yining Wang and Aarti Singh. **Provably Correct Active Sampling Algorithms for Matrix Column Subset Selection with Missing Data**. *arXiv:1505.04343*.

JOURNAL PUBLICATIONS

[3] Yan Xu, Yining Wang, Tianren Liu, Jiahua Liu, Yubo Fan, Yi Qian, Junichi Tsujii and Eric Chang. **Joint Segmentation and Named Entity Recognition using Dual Decomposition in Chinese discharge summaries.** *Journal of the American Medical Informatics Association* 21.e1:e84-e92, 2014.

[4] Yan Xu, Yining Wang, Jian-Tao Sun, Jianwen Zhang, Junichi Tsujii and Eric Chang. **Building Large Collections of Chinese and English Medical Terms from Semi-structured and Encyclopedia Websites.** *PLOS One* 8(7):e67526, 2013.

[5] Yan Xu, Yining Wang, Tianren Liu, Junichi Tsujii and Eric Chang. **An end-to-end system to identify temporal relation in discharge summaries: 2012 i2b2 challenge.** *Journal of the American Medical Informatics Association* 20:840-858, 2013.

[6] Yong Cui, Lian Wang, Xin Wang, Hongyi Wang and Yining Wang. **FMTCP: A Fountain Code-base Multipath Transmission Control Protocol.** *ACM/IEEE Transactions on Networking*, in press.

PEER-REVIEWED CONFERENCE PUBLICATIONS

[7] Yining Wang and Animashree Anandkumar. **Online and Differentially-Private Tensor Decomposition**. In *Advances in Neural Information Processing Systems (NIPS)*, 2016.

[8] Bo Li*, Yining Wang*, Aarti Singh and Yevgeniy Vorobeychik. **Data poisoning attacks on factorization-based collaborative filtering**. In *Advances in Neural Information Processing Systems (NIPS)*, 2016.

[9] Maria-Florina Balcan*, Simon Du*, Yining Wang* and Adams Wei Yu*. **An Improved Gap-Dependency Analysis of the Noisy Power Method**. In *Conference on Learning Theory (COLT)*, 2016.

[10] Yining Wang, Yu-Xiang Wang and Aarti Singh. **Graph Connectivity in Noisy Sparse Subspace Clsutering**. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2016.

[11] Yining Wang and Aarti Singh. **Noise-adaptive Margin-based Active Learning and Lower Bounds under Tsybakov Noise Condition**. In *AAAI Conference on Artificial Intelligence (AAAI)*, 2016 (selected for oral presentation).

[12] Yining Wang, Hsiao-Yu Tung, Animashree Anandkumar and Alexander Smola. **Fast and Guaranteed Tensor Decomposition**. In *Advances in Neural Information Processing Systems (NIPS)*, 2015 (selected for spotlight presentation, 67 of 1838 submissions).

[13] Yining Wang, Yu-Xiang Wang and Aarti Singh. **Differentially Private Subspace Clustering**. In *Advances in Neural Information Processing Systems (NIPS)*, 2015.

[14] Yining Wang, Yu-Xiang Wang and Aarti Singh. **A Deterministic Analysis of Noisy Sparse Subspace Clustering for Dimensionality-reduced Data.** In *International Conference on Machine Learning (ICML)*, 2015.

[15] Yining Wang and Jun Zhu. **DP-Space: Bayesian Nonparametric Subspace Clustering with Small-variance Asymptotics.** In *International Conference on Machine Learning (ICML)*, 2015.

[16] Yining Wang and Aarti Singh. **Column Subset Selection with Missing Data via Adaptive Sampling.** In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2015.

[17] Yining Wang and Jun Zhu. **Spectral Methods for Supervised Topic Models.** In *Advances in Neural Information Processing Systems (NIPS)*, 2014.

[18] Yining Wang and Jun Zhu. **Small-variance Asymptotics for Gibbs Infinite SVMs.** In *AAAI Conference on Artificial Intelligence (AAAI)*, 2014 (selected for oral presentation).

[19] Jiajun Wu, Yining Wang, Zhulin Li and Zhuowen Tu. **Harvesting Motion Patterns in Still Images from the Internet.** In *Annual Meeting of the Cognitive Science Society (CogSci)*, 2014 (selected for oral presentation).

[20] Yining Wang, Liwei Wang, Yuanzhi Li, Di He, Wei Chen and Tie-Yan Liu. **A Theoretical Analysis of Normalized Discounted Cumulative Gain (NDCG) Ranking Measures.** In *Conference on Learning Theory (COLT)*, 2013.

[21] Jingjing Liu, Panupong Pasupat, Yining Wang, Scott Cyphers and Jim Glass. **Query Understanding Enhanced by Hierarchical Parsing Structures.** In *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2013.

[22] Yong Cui, Xin Wang, Hongyi Wang, Guangjin Pan and Yining Wang. **FMTCP: A Fountain Code-base Multipath Transmission Control Protocol.** In *International Conference on Distributed Computing Systems (ICDCS)*, 2012.

LIGHTLY/NON REVIEWED PUBLICATIONS

[23] Yining Wang and Aarti Singh. **An Empirical Comparison of Sampling Techniques for Matrix Column Subset Selection**. In *Annual Allerton Conference on Communication, Control and Computing (Allerton)*, 2015.

RESEARCH EXPERIENCE

Microsoft Research Labs, Redmond, WA, USA

Research Intern **May 2016 to August 2016**

- Supervisor: Dengyong (Denny) Zhou.
- Design and implementation of novel recurrent neural network structures for machine translation and automatic speech recognition.

Symantec Research Labs, Culver City, CA, USA

Research Intern **June 2015 to August 2015**

- Supervisor: Petros Efstathopoulos and Kevin Roundy.
- Design and implementation of Project Harbinger, a system for enterprise level malicious attack prediction based on collaborative filtering.

Tsinghua University, Beijing, China

Research Assistant, Dept. of Computer Science and Technology **August 2013 to July 2014**

- Advisor: Prof. Jun Zhu
- Speeding up infinite SVM, a Bayesian nonparametric supervised model, by small-variance asymptotic analysis.
- Spectral methods for supervised latent variable models.

Massachusetts Institute of Technology, Cambridge, MA, USA

Undergraduate Research Opportunity Program (UROP) **February 2013 to June 2013**

- Advisor: Dr. Jingjing Liu and Prof. Jim Glass
- Improving semantic role labeling in Movie Browser (a spoken dialogue system) by incorporating hierarchical parsing and semantic dependency.
- Collected human labeled edX discussion threads for question answering via crowdsourcing on the Amazon Mechanical Turk platform.

Research Assistant **February 2013 to June 2013**

- Advisor: Prof. Cynthia Rudin
- Mixed integer optimization for learning to rank and rerank.

Microsoft Research Asia, Beijing, China

Research Intern, TechStrategy Group **October 2011 to January 2013**

- Supervisor: Dr. Eric Chang and Prof. Junichi Tsujii
- Participated in the 2012 I2b2 shared task. We developed an end-to-end system to extract event and time expressions from clinical records and then to classify temporal relations between entities. Our system ranked the first on 2 of the 3 tracks.
- Built large collections of multilingual medical terms by exploiting parallel structures in webpages and searching snippets.
- Annotation of Chinese clinical records. We also applied dual decomposition to improve the performance of both named entity recognition and word segmentation tasks.

AWARDS AND HONORS

Tsinghua University scholarships

- Yao Award (the highest honor for Yao Class students), second award (2/35). 2013
- ZHANG Mingwei Scholarship. 2013
- Baidu Future Star Scholarship. 2012
- Baidu Scholarship, second award. 2011
- Freshman Scholarship. 2010

Student travel scholarships
- AAAI 2014, NIPS 2014, NIPS 2015.

Programming contests
- Silver Medal. The 35th ACM/ICPC Regional Contest Chengdu Site. 2010
- Gold Medal. Chinese National Olympiad in Informatics. 2009

PROFESSIONAL SERVICE

Reviewer
- **Conferences**: ICML, NIPS, AISTATS, ECCV, COLT.
- **Journals**: Journal of Machine Learning Research, IEEE Transactions on Pattern Analysis and Machine Intelligence, Journal of the Royal Statistics Society Series C (Applied Statistics), IEEE Transactions on Signal Processing, Digital Signal Processing.

Conference volunteer
- ICML 2014, AAAI 2014.

SKILLS

Machine Learning and Numerical Algebra Packages
- Lapack, Arpack, Armadillo, Petuum, Theano, Torch

Programming Languages
- Java, C/C++, C#, Pascal, Python, Javascript, HTML5, SQL, MATLAB, LaTeX, Lua

Operating Systems
- Windows, Linux