

MINMIN CHEN

Department of Computer Science and Engineering
Washington University in St. Louis
Email: mc15@cse.wustl.edu
Phone: +1 832 867 3200

One Brookings Drive, Campus Box 1045
Saint Louis, MO 63130
URL: <http://www.cse.wustl.edu/~mchen>

EDUCATION

WASHINGTON UNIVERSITY IN ST. LOUIS *2006-present*

Ph.D. candidate in Dept. of Computer Science and Engineering

Thesis advisor: Prof. Kilian Q. Weinberger

Thesis title: Domain Adaptation through Learned Data Representations

Thesis committee: John Blitzer, John Cunningham, Tao Ju, Robert Pless, Bill Smart

WASHINGTON UNIVERSITY IN ST. LOUIS *2006-2009*

Master of Science in Dept. of Computer Science and Engineering

Research advisor: Prof. Yixin Chen

Project: Conditional Random Fields and its Application in Computational Genetics

GPA: 4.0/4.0

TSINGHUA UNIVERSITY *2002-2006*

Bachelor of Engineering in Dept. of Electronic and Communications

RESEARCH EXPERIENCE **Fields:** Machine learning, Optimization

Themes: My research focuses on two aspects in large scale learning: (1) How to make use of different levels of supervision carried in data to learn? [publication 16, 15, 13, 12, 9, 8, 7, W5, W4, W1] (2) How to develop algorithms that can handle data of such scale at runtime? [publication 14, 11, W3]

Topics: domain adaptation/transfer learning, learning with weak supervision, semi-supervised learning, budgeted learning, large scale sequential learning

Applications: Text mining, Image classification, Ranking, Personalization, Computational biology, Healthcare

INDUSTRIAL INTERNSHIP YAHOO! LABS

Fall 2012

Research intern in Advertising Science team

Manager: Prabhakar Krishnamurthy

User behavior targeting: reinventing the training pipeline for matching online advertisement to potential clickers based on user behavior history

Coding with perl, bash script, and hadoop map-reduce

MICROSOFT RESEARCH, REDMOND

Summer 2012

Research intern in Machine Learning department

Mentor: Alice Zheng

Machine aided labeling: reducing complex labeling tasks, common in large scale learning, to simpler ones through the use of learning algorithms. Developing an automatic image annotation system that learns from incomplete user tags.

Coding with matlab, experimenting Amazon Mechanical Turk.

MICROSOFT RESEARCH, ASIA (BEIJING)

Summer 2010

Research intern in Machine Learning department

Mentor: Jian-Tao Sun

Context-aware query classification: improving query classification by exploiting search contexts embedded in query sessions and unlabeled queries free on the web.

Coding with C/C++.

PROFESSIONAL SERVICE **Organizer**

-- the 7th Women in Machine Learning workshop, Lake Tahoe Nevada, 2012

Program committee member

-- International Conference on Machine Learning (ICML)

-- Advances in Neural Information Processing Systems (NIPS)

-- AAAI Conference on Artificial Intelligence (AAAI)

-- Conference on Uncertainty in Artificial Intelligence (UAI)

Reviewer

-- Journal of Machine Learning

AWARD

Award of Excellence

-- Microsoft Stars of Tomorrow Internship program

Finalist of Best Paper Award

-- 15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining

TEACHING

WASHINGTON UNIVERSITY IN ST LOUIS

Teaching assistant for

-- Machine Learning (CSE 717A, spring 2010)

-- Algorithm for Nonlinear Optimization (CSE 543T, spring 2010)

-- Artificial Intelligence (CSE 511, fall 2009)

TALKS

Automatic Image Annotation using Incomplete User Tags, Microsoft Research, Redmond, 2012.

Marginalized Stacked Denoising Autoencoders for Domain Adaptation, ICML, Edinburgh, 2012

Classifier Cascade for Minimizing Feature Evaluation Cost, AISTATS, Canary island, 2012

Marginalized Stacked Denoising Autoencoders, Snowbird, Utah, 2012

Co-training for Domain Adaptation, NIPS, Spain, 2011

Classifier Cascade: Tradeoff Accuracy and Feature Evaluation Cost, WiML, Spain, 2011

Improving Context-Aware Query Classification via Adaptive Self-Training, CIKM, England, 2012

Automatic Feature Decomposition for Single View Co-training, ICML, Seattle, 2011

Constrained Optimization for Validation-Guided CRFs, KDD, France, 2009

CRF-OPT: An Efficient High-Quality Condition Random Field Solver, AAAI, Chicago, 2008

SELECTED PUBLICATIONS L. Maaten, M. Chen, K. Weinberger, S. Tyree, **Learning from Marginalized Corrupted Features**, *30th International Conference on Machine Learning (ICML)*, 2013.

M. Chen, Z. Xu, K. Weinberger, F. Sha, **Marginalized Stacked Denoising Auto-encoders for Domain Adaptation**, *29th International Conference on Machine Learning (ICML)*, 2012.

M. Chen, K. Weinberger, O. Chapelle, Z. Xu, D. Kedem, **Classifier Cascade for Minimizing Feature Evaluation Cost**, *15th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2012. *Oral presentation (26/400)*.

M. Chen, K. Weinberger, J. Blitzer, **Co-training for Domain Adaptation**, *25th Annual Conference on Neural Information Processing Systems (NIPS)*, 2011.

M. Chen, K. Weinberger, Y. Chen, **Automatic Feature Decomposition for Single View Co-training**, *28th International Conference on Machine Learning (ICML)*, 2011.

REFERENCES

Available upon request

PUBLICATIONS

- [16] M. Chen, K. Weinberger, A. Zheng, **Automatic Image Annotation using Incomplete User Tags**, *under preparation*.
- [15] L. Maaten, M. Chen, K. Weinberger, S. Tyree, **Learning from Marginalized Corrupted Features**, *30th International Conference on Machine Learning (ICML), 2013*.
- [14] Z. Xu, M. Kusner, K. Weinberger, M. Chen, **Cost-Sensitive Trees of Classifiers**, *30th International Conference on Machine Learning (ICML), 2013*.
- [13] Z. Xu, M. Chen, K. Weinberger, F. Sha, **From sBoW to dCoT: Marginalized Encoders for Text Representation**, *21st ACM Conference on Information and Knowledge Management (CIKM), 2012*.
- [12] M. Chen, Z. Xu, K. Weinberger, F. Sha, **Marginalized Stacked Denoising Auto-encoders for Domain Adaptation**, *29th International Conference on Machine Learning (ICML), 2012*.
- [11] M. Chen, K. Weinberger, O. Chapelle, Z. Xu, D. Kiedem, **Classifier Cascade for Minimizing Feature Evaluation Cost**, *15th International Conference on Artificial Intelligence and Statistics (AISTATS), 2012*.
- [10] Y. Mao, Y. Chen, G. Hackmann, M. Chen, C. Lu, M. Kollef, and T. Bailey, **Early Deterioration Warning for Hospitalized Patients by Mining Clinical Data**, *International Journal of Knowledge Discovery in Bioinformatics, 2(3):1-20, 2012*.
- [9] M. Chen, K. Weinberger, J. Blitzer, **Co-training for Domain Adaptation**, *25th Annual Conference on Neural Information Processing Systems (NIPS), 2011*.
- [8] M. Chen, K. Weinberger, Y. Chen, **Automatic Feature Decomposition for Single View Co-training**, *28th International Conference on Machine Learning (ICML), 2011*.
- [7] M. Chen, J. Sun, X. Ni, Y. Chen, **Improving Context-Aware Query Classification via Adaptive Self-training**, *20th ACM Conference on Information and Knowledge Management (CIKM), 2011*.
- [6] G. Hackmann, M. Chen, O. Chipara, C. Lu, Y. Chen, M. Kollef, and T. Bailey, **Toward a Two-Tier Clinical Warning System for Hospitalized Patients**, *American Medical Informatics Association Annual Symposium (AMIA), 2011*.
- [5] Y. Chen, M. Chen, **Extended Duality for Nonlinear Programming**, *Journal of Computational Optimization and Applications, 2010*.

[4] M. Chen, **To Improve the Speed and Generalization Ability of Conditional Random Fields**, *Master's Thesis, Dept. of Computer Science and Engineering, Washington University in St. Louis, 2009.*

[3] M. Chen, Y. Chen, M. Brent, A. Tenney, **Constrained Optimization for Validation-Guided Conditional Random Field Learning**, *15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2009.*

[2] M. Chen, Y. Chen, M. Brent, A. Tenney, **Gradient-Based Feature Selection for Conditional Random Fields and Its Applications in Computational Genetics**, *21st IEEE International Conference on Tools with Artificial Intelligence (ICTAI), 2009.*

[1] M. Chen, Y. Chen, M. Brent, **CRF-OPT: An Efficient High-Quality Conditional Random Field Solver**, *23rd Conference on Artificial Intelligence (AAAI), 2008.*

WORKSHOPS

[W5] M. Chen, K. Weinberger, A. Zheng, **Learning from Incomplete User Tags**, *NIPS workshop on Large Scale Visual Recognition and Retrieval, 2012.*

[W4] M. Chen, Z. Xu, K. Weinberger, F. Sha, **Marginalized Stacked Denoising Auto-encoders**, *The Learning Workshop (Snowbird), 2012.*

[W3] M. Chen, K. Weinberger, O. Chapelle, **Classifier Cascade: Tradeoff between Accuracy and Feature Evaluation Cost**, *6th Annual Workshop for Women in Machine Learning (WiML), 2011.*

[W2] Y. Mao, Y. Chen, G. Hackmann, M. Chen, C. Lu, M. Kollef, and T. Bailey, **Medical Data Mining for Early Deterioration Warning in General Hospital Wards**, *ICDM workshop on Biological Data Mining and its Applications in Healthcare(BioDM), 2011.*

[W1] M. Chen, J. Sun, X. Ni, **Adaptive Self-Training with Max Margin Conditional Random Fields for Context Aware Query Classification**, *5th Annual workshop for Women in Machine Learning (WiML), 2010.*