

Kai-Wei Chang

University of California, Los Angeles
 4732 Boelter Hall
 Los Angeles, CA 90095

kw@kwchang.net
<http://kwchang.net>

RESEARCH INTERESTS

- *Statistical approaches to natural language processing.*
- *Tractable machine learning methods for complex and big data.*

EDUCATION AND EXPERIENCE**University of California Los Angeles, CA**

Assistant Professor, Computer Science

Aug. 2017 –

University of Virginia, VA

Assistant Professor, Computer Science

Aug. 2016 – Jun. 2017

Microsoft, MA

Postdoctoral researcher, Microsoft Research New England Lab

Jul. 2015 – Aug. 2016

University of Illinois at Urbana-Champaign, IL

Ph.D. in Computer Science

Aug. 2010 – May. 2015

Cognitive Computation Group; w/ Dan Roth

National Taiwan University, Taipei, Taiwan

M.S. in Computer Science and Information Engineering

Jun. 2009

Machine Learning and Data Mining Group; w/ Chih-Jen Lin

B.S. in Computer Science and Information Engineering

Jun. 2007

B.S. in Electrical Engineering (Dual Degree)

Jun. 2007

Microsoft

Intern, Microsoft Research, New York (Mentor: John Langford, publications [6,39,40])

2014

Intern, Microsoft Research, Redmond (Mentor: Scott Wen-tau Yih, publications [12,9])

2013

Intern, Microsoft Silicon Valley (Mentor: Sathya Keerthi, publication [14])

2012

Google

Intern, Google Research Group, Beijing (Advisor: Chih-Jen Lin, publications [24,43])

2008

SELECTED AWARDS

Best Long Paper Award, EMNLP 2017

2017

Best Research Paper Award, SIGKDD 2010

2010

Yahoo! Key Science Challenge Award, Yahoo!

2011

C. L. and Jane W. S. Liu Award, University of Illinois

2013

Given in support of a student showing exceptional research promise relatively early in their graduate studies.

Facebook PhD Fellowship Finalist

2014

Microsoft Research PhD Fellowship Finalist

2013

Studying Abroad Scholarship, Ministry of Education, Taiwan

2009

PUBLICATIONS

7,300+ Google Scholar citations in total.

Published at ICML, NIPS, EMNLP, AAAI, ACL, JMLR, KDD, TKDD, ICDM, ECML, CoNLL

Refereed Conference Publications

- [1] J. Zhao, T. Wang, M. Yatskar, V. Ordonez, **K.-W. Chang**, Men Also Like Shopping: Reducing Gender Bias Amplification using Corpus-level Constraints. *Conference on Empirical Methods in Natural Language Processing (EMNLP 2017)*, **Best Long Paper Award**.
- [2] T. Bolukbasi, **K.-W. Chang**, Joseph Wang, Venkatesh Saligrama. Structured Prediction with Test-time Budget Constraints. *Thirty-First AAAI Conference on Artificial Intelligence (AAAI 2017)*.
- [3] T. Bolukbasi, **K.-W. Chang**, James Zou, Venkatesh Saligrama, Adam Kalai, Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. *Neural Information Processing Systems (NIPS 2016)*.
- [4] K.-W. Chang, H. He, H. Daume III, J. Langford, S. Ross A Credit Assignment Compiler for Joint Prediction. *Neural Information Processing Systems (NIPS 2016)*.
- [5] S. Upadhyay, M. Chang, **K.-W. Chang**, W.-t. Yih, Learning from Explicit and Implicit Supervision Jointly For Algebra Word Problems, *Conference on Empirical Methods in Natural Language Processing (EMNLP 2016)*.
- [6] **K.-W. Chang**, A. Krishnamurthy, A. Agarwal, H. Daumé III, J. Langford. Learning to search better than your teacher, *the 32nd International Conference on Machine Learning (ICML 2015)*.
- [7] H. Peng, **K.-W. Chang**, D. Roth. A joint framework for coreference resolution and mention head detection, *The SIGNLL Conference on Computational Natural Language Learning (CoNLL 2015)*.
- [8] **K.-W. Chang**, S. Upadhyay, G. Kundu and D. Roth Structural learning with amortized inference *The Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI 2015)*.
- [9] **K.-W. Chang**, W.-t. Yih, B. Yang and C. Meek. Typed tensor decomposition of knowledge bases for relation extraction. *Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*.
- [10] R. Samdani, **K.-W. Chang**, D. Roth. A discriminative latent variable model for online clustering. *the 31st International Conference on Machine Learning (ICML 2014)*.
- [11] **K.-W. Chang**, R. Samdani, D. Roth. A constrained latent variable model for coreference resolution. *Conference on Empirical Methods in Natural Language Processing (EMNLP 2013)*.
- [12] **K.-W. Chang**, W.-t. Yih, C. Meek. Multi-relational latent semantic analysis. *Conference on Empirical Methods in Natural Language Processing (EMNLP 2013)*.
- [13] **K.-W. Chang**, V. Srikumar, D. Roth. Multi-core structural SVM training. *European Conference on Machine Learning (ECML 2013)*.
- [14] **K.-W. Chang**, S. Sundararajan, S. S. Keerthi. Tractable semi-supervised learning of complex structured prediction models. *European Conference on Machine Learning (ECML 2013)*.
- [15] **K.-W. Chang**, B. Deka, W.-M. H. Hwu, D. Roth. Efficient pattern-based time series classification on GPU. *2012 IEEE 12th International Conference on Data Mining (ICDM 2012)*.
- [16] **K.-W. Chang** and D. Roth, Selective block minimization for faster convergence of limited memory large-scale linear models. *the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2011)*.
- [17] H.-F. Yu, C.-J. Hsieh, **K.-W. Chang**, and C.-J. Lin, Large linear classification when data cannot fit in memory, *the 22nd International Joint Conferences on Artificial Intelligence (IJCAI 2011, the Best Paper Track)*.

- [18] H.-F. Yu, C.-J. Hsieh, **K.-W. Chang**, and C.-J. Lin, Large linear classification when data cannot fit in memory, *the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2010)*, **Best research paper**.
- [19] F.-L. Huang, C.-J. Hsieh, **K.-W. Chang**, and C.-J. Lin, Iterative scaling and coordinate descent methods for maximum entropy models, *the 47th Annual Meeting of the Association for Computational Linguistics (ACL 2009, short paper)*.
- [20] S. S. Keerthi, S. Sundararajan, **K.-W. Chang**, C.-J. Hsieh, and C.-J. Lin, A sequential dual method for large scale multi-class linear SVMs, *the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2008)*.
- [21] C.-J. Hsieh, **K.-W. Chang**, C.-J. Lin, S. Sathya Keerthi, and S. Sundararajan, A dual coordinate descent method for large-scale linear SVM, *the 25th International Conference on Machine Learning (ICML 2008)*.

Refereed Journal Publications

- [22] H.-F. Yu, C.-J. Hsieh, **K.-W. Chang**, and C.-J. Lin, Large linear classification when data cannot fit in memory, *ACM Transactions on Knowledge Discovery from Data (TKDD)* 5(4):23, 2012
- [23] G.-X. Yuan, **K.-W. Chang**, C.-J. Hsieh, C.-J. Lin, A comparison of optimization methods for large-scale L1-regularized linear classification. *Journal of Machine Learning Research 11 (JMLR)*, 3183-3234, 2010.
- [24] Y.-W. Chang, C.-J. Hsieh, **K.-W. Chang**, Michael Ringgaard, and C.-J. Lin, Training and Testing Low-degree Polynomial Data Mappings via Linear SVM *Journal of Machine Learning Research 11 (JMLR)*, 1471-1490, 2010.
- [25] F.-L. Huang, C.-J. Hsieh, **K.-W. Chang**, and C.-J. Lin, Iterative scaling and coordinate descent methods for maximum entropy models. *Journal of Machine Learning Research 11 (JMLR)*, 815-848, 2010.
- [26] R.-E. Fan, **K.-W. Chang**, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research 9 (JMLR)*, 1871-1874, 2008.
- [27] **K.-W. Chang**, C.-J. Hsieh, and C.-J. Lin, Coordinate descent method for large-scale L2-loss linear SVM. *Journal of Machine Learning Research 9 (JMLR)*, 1369-1398, 2008.

Refereed Shared Task System Papers

- [28] A. Rozovskaya, **K.-W. Chang**, D. Roth. The Illinois-Columbia System in the CoNLL-2014 Shared Task *Proceedings of the Eighteenth Conference on Computational Natural Language Learning (CoNLL 2014)* (**1st place**).
- [29] A. Rozovskaya, **K.-W. Chang**, M. Sammons, D. Roth. The University of Illinois System in the CoNLL-2013 Shared Task. *Proceedings of the Seventeenth Conference on Computational Natural Language Learning (CoNLL) 2013* (**1st place**).
- [30] X. Cheng, B. Chen, R. Samdani, **K.-W. Chang**, Z. Fei, M. Sammons, J. Wieting, S. Roy, C. Wang, and D. Roth, Illinois Cognitive Computation Group UI-CCG TAC 2013 Entity Linking and Slot Filler Validation Systems, *Text Analysis Conference (TAC 2013)*.
- [31] **K.-W. Chang**, R. Samdani, A. Rozovskaya, M. Sammons and D. Roth, Illinois-Coref: The UI System in the CoNLL-2012 Shared Task *Proceedings of the Sixteenth Conference on Computational Natural Language Learning (CoNLL) 2012* (**4th place in the English closed track**).
- [32] **K.-W. Chang**, R. Samdani, A. Rozovskaya, N. Rizzolo, M. Sammons and D. Roth, Inference Protocols for Coreference Resolution. *Proceedings of the Fifteenth Conference on Computational Natural Language Learning (CoNLL) 2011* (**3rd place in the closed track**).

- [33] H.-Y. Lo, **K.-W. Chang**, S.-T. Chen, T.-H. Chiang, C.-S. Ferng, C.-J. Hsieh, Y.-K. Ko, T.-T. Kuo, H.-C. Lai, K.-Y. Lin, C.-H. Wang, H.-F. Yu, C.-J. Lin, H.-T. Lin and S.-D. Lin. An ensemble of three classifiers for KDD Cup 2009: expanded linear model, heterogeneous boosting, and selective naive Bayes. *Proceedings of KDD-Cup 2009 competition, vol. 7 of JMLR Workshop and Conference Proceedings*, 57-64, 2009. **(3rd Place out of 400+ submissions in the Slow Track).**

Thesis

- [34] K.-W. Chang, Selective algorithms for large-scale classification and structured learning (Ph.D.)
 [35] K.-W. Chang, A dual coordinate descent method for large-scale linear SVM (Master)

Workshop Publications and Preprints

- [36] S. Upadhyay, **K.-W. Chang**, M. Taddy, A. Kalai, J. Zou. Beyond Bilingual: Multi-senseWord Embeddings using Multilingual Context. *ACL Representation learning for NLP Workshop 2017*, **Best Paper Award**.
- [37] K. Arnold, **K.-W. Chang**, A Kalai. Learning to Suggest Phrases. *AAAI Workshop on Human-Aware AI Workshop*, 2017.
- [38] C.-p. Lee, K.-W. Chang, S. Upadhyay, D. Roth. Distributed Training of Structured SVM. *NIPS Workshop on Optimization for Machine Learning*, 2015.
- [39] K. -W. Chang, H. Daumé III, J. Langford, S. Ross. Efficient Programmable Learning to Search. *ICML Workshop on Machine Learning System*, 2015.
- [40] K.-W. Chang, H. He, H. Daumé III, J. Langford. Learning to Search for Dependencies. *Arxiv 1503.05615*, 2015.
- [41] R. Samdani, **K.-W. Chang**, D. Roth. A Discriminative Latent Variable Model for Clustering of Streaming Data with Application to Coreference Resolution. *ICML workshop on Inference: Interactions between Inference and Learning*, 2013.
- [42] H.-F. Yu, C.-J. Hsieh, **K.-W. Chang**, and C.-J. Lin, Pascal Challenge: Linear Support Vector Machines. *Pascal Large Scale Learning Challenge in ICML 2008 Workshop*, 2008.

Patents

- [43] Efficient polynomial mapping of data for use with linear support vector machines, Y.-W. Chang, C.-J. Hsieh, **K.-W. Chang**, M. Ringgaard, C.-J. Lin, 2013.
- [44] Interactive Context-Based Text Completions, K. Arnold, **K.-W. Chang**, A Kalai, 2016 (under review).

SOFTWARE DESIGN

- | | |
|---|-----------------------|
| Illinois Coreference Resolution Demo | Jan. 2013 |
| http://cogcomp.cs.illinois.edu/demo/corefdemo | |
| A Library for Learning Linear Model with Limited Memory | Sep. 2011 |
| http://cogcomp.cs.illinois.edu/page/software_view/LMLM | |
| LIBLINEAR - A Library for Large Linear Classification | Sep. 2007 |
| http://www.csie.ntu.edu.tw/~cjlin/liblinear | |
| - One of the main contributors. | |
| - The library has been used in some major Internet companies to classify their web data. | |
| - More than 60,000 downloads since Apr. 2007. | |
| LIBSVM - A Library for Support Vector Machine | Sep. 2006 – Aug. 2010 |
| http://www.csie.ntu.edu.tw/~cjlin/libsvm | |

- Assisted Prof. Chih-Jen Lin in maintaining the library and answering questions from users.
- The library has been downloaded more than 300,000 times since Apr. 2000.

TEACHING EXPERIENCE

University of Virginia, VA

- Instructor, Advanced Machine Learning, Spring 2017.
- **Instructor, Natural Language Processing, Fall 2016.**

University of Illinois, IL

- Head Teaching Assistant, Machine Learning, Dan Roth, Fall 2014 (150+ registered students).
- Guest Instructor (three lectures), Machine Learning, Dan Roth, Fall 2014.

National Taiwan University, Taipei, Taiwan

- Teaching Assistant, Formal Languages and Automata, Chih-Jen Lin, Fall 2008.
- Teaching Assistant, Numerical Methods, Chih-Jen Lin, Spring 2009.

TALKS

Tutorials

- **Learning and Inference in Structured Prediction Models, K.W. Chang, G. Kundu, D. Roth, and V. Srikumar. AAI 2016.**
- Hands-on Learning to Search for Structured Prediction, H. Daumé III, J. Langford, K.-W. Chang, H. He, and S. Rao. NAACL 2015.

Invited Talks

- University of Utah, Aug 2017.
- Appier, June 2017.
- National Taiwan University, June 2017.
- University of California, Los Angeles, April 2017.
- University of Massachusetts Amherst, March 2017.
- University of California, Davis, March 2017.
- NIPS workshop on learning high dimensions with structure, Dec 2016.
- University of Massachusetts Medical School, May 2016.
- University of Massachusetts Amherst, November 2015.
- Toyota Technological Institute at Chicago, April 2015.
- University of Virginia, March 2015.
- Boston University, March 2015.
- The Ohio State University, March 2015.
- Carnegie Mellon University, March 2015.
- Washington State University, March 2015.
- Microsoft Research, Redmond, February 2015.
- University of Arizona, February 2015.
- Mid-Atlantic Student Colloquium on Speech, Language and Learning, Breakout Session, Johns Hopkins University, January 2015.
- Microsoft Research, Cambridge, MA, January 2015.
- Purdue University, Machine Learning Seminar, December 2014.

- University of Maryland, Seminar Talk, November 2014.
- Columbia University, NLP Seminar, November 2014.
- University of Illinois, Data&Information System Seminar, November 2014.
- Microsoft Research, New York, July 2014.
- University of Illinois, Data&Information System Seminar, November 2013.
- Microsoft Research, Redmond, July 2012.

Conference Talks

- ICML 2014: A Discriminative Latent Variable Model for Online Clustering.
- EMNLP 2013: Multi-Relational Latent Semantic Analysis.
- ECML 2013: Multi-core Structural SVM Training.
- ECML 2013: Tractable Semi-Supervised Learning of Complex Structured Prediction Models.
- ICDM 2012: Efficient Pattern-Based Time Series Classification on GPU.
- KDD 2011: Selective Block Minimization for Faster Convergence of Limited Memory Large-Scale Linear Models.
- CoNLL 2011: Inference Protocols for Coreference Resolution.
- IJCAI 2011: Large Linear Classification when Data cannot Fit in Memory.
- ICML 2008: A Dual Coordinate Descent Method for Large-Scale Linear SVM.

PROFESSIONAL ACTIVITIES

Organizer:

- Workshop on Deep Structured Prediction. Co-organizers: Isabelle Augenstein, Kai-Wei Chang, Gal Chechik, Bert Huang, Andre Martins, Ofer Meshi, Yishu Miao, Alexander Schwing. ICML 2017.
- 2nd Workshop on Structured Prediction for Natural Language Processing. Co-organizers: K.-W. Chang, M. Chang, A. Rush, and V. Srikumar. EMNLP 2017.
- Structured Prediction for Natural Language Processing. Co-organizers: K.-W. Chang, M. Chang, A. Rush, and V. Srikumar. EMNLP 2016.

Area Chair: NAACL 2018.

Program Committee/Reviewer:

- Machine Learning: ICML, ICLR, NIPS, JMLR, TNN, TNNLS, Neurocomputing, Neural Computation.
- Artificial Intelligence: AAI, WCCI, IJRAI, AISTATS.
- Natural Language Processing: ACL, CoNLL, Coling, TACL.
- Data Mining: KDD.

Internal Service. Department of Computer science, University of Illinois

- Student Volunteer in the Mentee-Mentor Program for new graduate students. 13', 14', 15'.
- Student Ambassador for assisting prospective students. 13', 14'.
- Student Volunteer for graduate admission. 13'.

OTHER HONORS

Research:

- Master Thesis Award, Taiwanese Association for Artificial Intelligence 2009
- Scholarship for Graduate Student, GARMIN 2008
- Honorary Member of the Phi Tau Phi Scholastic Honor Society, National Taiwan University 2009
Given to top 3% master students.
- Undergraduate Research Grant, National Science Council, Taiwan 2006

Shared Tasks:

- First Place in CoNLL-shared Task 2014 2014
 - First Place in CoNLL-shared Task 2013 2013
 - Fourth Place in CoNLL-shared Task 2012 English track 2012
 - Third Place in CoNLL-shared Task 2011 2011
- CoNLL shared task is the most prestigious competition in NLP.**
- Third Prize in the Slow Track of KDDCUP 2009 2009
Out of 400+ submissions. KDDCup is the most prestigious competition in data mining.

Academy:

- Presidential Award, National Taiwan University Fall 2003, Spring 2004, Fall 2004, Spring 2005, Fall 2006
Given to the top 5% undergraduate students each semester.
- Dean's List, National Taiwan University 2006
Given to top ranked students who provide academic consulting.

REFERENCES**Prof. Dan Roth**

Professor, Department of Computer Science
University of Illinois at Urbana-Champaign

*3322 Siebel Center
201 N. Goodwin Avenue
Urbana, IL 61801
Phone: (217) 244-7068
Fax: (217) 265-6591
Email: danr@illinois.edu*

Prof. Chih-Jen Lin

Distinguished Professor, Department of Computer Science and Information Engineering,
National Taiwan University

*Room 413, Department of Computer Science
National Taiwan University
Taipei, Taiwan, 106
Phone: (886) 2-2362-5336 ext. 413
Fax: (886) 2-2362-8167
Email: cjlin@csie.ntu.edu.tw*

Dr. Adam Kalai

Principle Researcher, Microsoft Research, Cambridge, MA

*Microsoft Research, Office 14063
One Memorial Drive
Cambridge, MA 02142
Phone: (857) 453-6323
Fax: (425) 936-7329
Email: adum@microsoft.com*

Dr. Scott Wen-tau Yih

Senior Researcher, Microsoft Research, Redmond, WA

*1 Microsoft Way
Redmond, WA 98052
Phone: (425) 706-0579
Fax: (425) 706-7329
Email: scottyih@microsoft.com*