Yuchen Zhang

Email: zhangyuc@gmail.com **Phone**: (+1)-510-423-1353

Education

University of California, Berkeley **2011 - 2016**
Ph.D. in Computer Science
Advised by Michael I. Jordan and Martin J. Wainwright
Research: distributed machine learning, optimization algorithms, deep learning theory

University of California, Berkeley **2011 - 2013**
Master in Statistics

Tsinghua University **2007 - 2011**
Bachelor in Computer Science
Supervised by Andrew C. Yao (Yao Class)

Employment History

Senior Researcher at Microsoft **2018 - Now**

- My team builds a revolutionary dialogue system that understands complex long-tail queries, and seamlessly drives multi-turn conversations to communicate, collaborate, and accomplish tasks. Serving MS Office and 3rd-party parteners with AI assiatant and customer service bots (demo link).
- I lead the core representation and modeling efforts for our dialogue system. I am the main designer and developer of the core dialogue representation, and a major contributor to the system's capability of compositional semantic parsing, state tracking, error handling and natural language generation.

Senior Research Scientist at Semantic Machines **2018**
Semantic Machines is a team of world-class researchers and engineers working on the next-generation dialogue system technologies. We were acquired to continue our mission at Microsoft since June, 2018.

Post-doc Researcher at Stanford University **2016 - 2018**
Research: question answering, semantic parsing, deep learning theory

Internships **2010 - 2016**
Microsoft Research, Google, Baidu

Research Projects

Natural language processing

- Task-oriented dialogue systems [J1, P1, P2].
- Semantic parsing for question answering [C2].

Machine learning and optimization

- Deep learning and non-convex optimization [C3,C4,C5,C6,C7].
- Convex optimization [J2,C10].
- Crowdsourcing [J6,C11].
- Robust machine learning [C1].
- Personalized recommender systems [C13].
- Web search and online advertising [C17,C18,C19,C20,C21,C22]

Distributed computing

- Distributed algorithms for machine learning [J7,J8,C8,C9,C15,C16].

- Foundamental theory of distributed computing [C12,C14].
- Programming interface for parallelizing stochastic algorithms [T1].

Other projects

- Theoretical statistics [J3,J4,J5].
- Theoretical computer science [J9,C23].

Patents

[P1] **Y. Zhang**, J. Wolfe, A. Pauls and D. Hall. Updating Constraints For Computerized Assistant Actions. *U.S. Patent*, Filed by Microsoft on March 30, 2020.

[P2] J. Andreas, A. Vorobev, A. Guo, J. Krishnamurthy, J. Bufe, J. Rusak and **Y. Zhang**. Response Generation For Conversational Computing Interface, *U.S. Patent*, Filed by Microsoft on August 12, 2019.

Journal Publications

[J1] **Y. Zhang** and the Semantic Machines team (alpha-beta order). Task-Oriented Dialogue as Dataflow Synthesis. *Transactions of the Association for Computational Linguistics*.

[J2] **Y. Zhang** and L. Xiao. Stochastic Primal-Dual Coordinate Method for Regularized Empirical Risk Minimization. *Journal of Machine Learning Research*.

[J3] X. Chen, A. Guntuboyina and **Y. Zhang**. A note on the approximate admissibility of regularized estimators in the Gaussian sequence model. *Electronic Journal of Statistics*.

[J4] **Y. Zhang**, M. Wainwright and MI. Jordan. Optimal prediction for sparse linear models? Lower bounds for coordinate-separable M-estimators. *Electronic Journal of Statistics*.

[J5] X. Chen, A. Guntuboyina and **Y. Zhang** (alpha-beta order). On Bayes Risk Lower Bounds. *Journal of Machine Learning Research*.

[J6] **Y. Zhang**, X. Chen, D. Zhou and MI. Jordan. Spectral Methods meet EM: A Provably Optimal Algorithm for Crowdsourcing. *Journal of Machine Learning Research*.

[J7] **Y. Zhang**, J. Duchi and M. Wainwright. Divide and Conquer Kernel Ridge Regression: A Distributed Algorithm with Minimax Optimal Rates. *Journal of Machine Learning Research*.

[J8] **Y. Zhang**, J. Duchi and M. Wainwright. Communication-Efficient Algorithms for Statistical Optimization. *Journal of Machine Learning Research*.

[J9] **Y. Zhang** and X. Sun. The Antimagicness of the Cartesian Product of Graphs. *Theoretical Computer Science*.

Conference Publications

[C1] **Y. Zhang**, P. Liang. Defending Against Whitebox Adversarial Attacks via Randomized Discretization. *Artificial Intelligence and Statistics (AISTATS)*, 2019.

[C2] **Y. Zhang**, P. Pasupat, P. Liang. Macro Grammars and Holistic Triggering for Efficient Semantic Parsing. *Empirical Methods on Natural Language Processing (EMNLP)*, 2017.

[C3] **Y. Zhang**, P. Liang, M. Wainwright. Convexified Convolutional Neural Networks. *International Conference on Machine Learning (ICML)*, 2017.

[C4] **Y. Zhang**, P. Liang, M. Charikar. A Hitting Time Analysis of Stochastic Gradient Langevin Dynamics. *Conference on Learning Theory (COLT)*, 2017 **(Best paper award)**.

[C5] **Y. Zhang**, JD. Lee, M. Wainwright and MI. Jordan. On the Learnability of Fully-connected Neural Networks. *Artificial Intelligence and Statistics (AISTATS)*, 2017.

[C6] C. Jin, **Y. Zhang**, S. Balakrishnan, M. Wainwright, MI. Jordan. Local Maxima in the Likelihood of Gaussian Mixture Models: Structural Results and Algorithmic Consequences. *Neural Information Processing Systems (NIPS)*, 2016.

[C7] **Y. Zhang**, JD. Lee, MI. Jordan. ℓ_1-regularized Neural Networks are Improperly Learnable in Polynomial Time. *International Conference on Machine Learning (ICML)*, 2016.

[C8] **Y. Zhang**, M. Wainwright and MI. Jordan. Distributed Estimation of Generalized Matrix Rank: Efficient Algorithms and Lower Bounds. *International Conference on Machine Learning (ICML)*, 2015.

[C9] **Y. Zhang** and L. Xiao. DiSCO: Communication-Efficient Distributed Optimization of Self-Concordant Loss. *International Conference on Machine Learning (ICML)*, 2015.

[C10] **Y. Zhang** and L. Xiao. Stochastic Primal-Dual Coordinate Method for Regularized Empirical Risk Minimization. *International Conference on Machine Learning (ICML)*, 2015.

[C11] **Y. Zhang**, X. Chen, D. Zhou and MI. Jordan. Spectral Methods meet EM: A Provably Optimal Algorithm for Crowdsourcing. *Neural Information Processing Systems (NIPS)*, 2014. **(Spotlight presentation, 4.8% acceptance rate)**

[C12] **Y. Zhang**, M. Wainwright and MI. Jordan. Lower Bounds on the Performance of Polynomial-time Algorithms for Sparse Linear Regression. *Conference on Learning Theory (COLT)*, 2014.

[C13] **Y. Zhang**, A. Ahmed, V. Josifovski and A. Smola. Taxonomy Discovery for Personalized Recommendation. *ACM International Conference on Web Search and Data Mining (WSDM)*, 2014.

[C14] **Y. Zhang**, J. Duchi, M. Wainwright and MI. Jordan. Information-theoretic Lower Bounds for Distributed Statistical Estimation with Communication Constraints. *Neural Information Processing Systems (NIPS)*, 2013. **(Oral presentation, 1.4% acceptance rate)**

[C15] **Y. Zhang**, J. Duchi and M. Wainwright. Divide and Conquer Kernel Ridge Regression. *Conference on Learning Theory (COLT)*, 2013.

[C16] **Y. Zhang**, J. Duchi and M. Wainwright. Communication-Efficient Algorithms for Statistical Optimization. *Neural Information Processing Systems (NIPS)*, 2012.

[C17] W. Chen, D. Wang, **Y. Zhang** and Q. Yang. Understanding Click Noise: A Noise-aware Click Model for Web Search. *ACM International Conference on Web Search and Data Mining (WSDM)*, 2012.

[C18] **Y. Zhang**, W, Chen and D, Wang, Q. Yang. User-click Modeling for Understanding and Predicting Search-behavior. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, 2011.

[C19] B. Hu, **Y. Zhang**, G. Wang, Q. Yang, W. Chen. Characterize Search Intent Diversity into Click Models. *International World Wide Web Conference (WWW)*, 2011.

[C20] **Y. Zhang**, D. Wang, G. Wang, W. Chen, Z. Zhang, B. Hu and L. Zhang. Learning Click Model via Probit Bayesian Inference. *ACM International Conference on Information and Knowledge Management (CIKM)*, 2010.

[C21] D. Wang, W. Chen, G. Wang, **Y Zhang** and B. Hu. Explore Click Models for Search Ranking. *ACM International Conference on Information and Knowledge Management (CIKM)*, short paper, 2010.

[C22] F. Zhong, D. Wang, G. Wang, W. Chen, **Y. Zhang**, Z. Chen and H. Wang. Incorporating Post-Click Behaviors Into a Click Model. *Annual International ACM SIGIR Conference (SIGIR)*, 2010.

[C23] **Y. Zhang** and L. Zhang. Extracting Independent Rules: a New Perspective of Boosting. *International Symposium on Artificial Intelligence and Mathematics (ISAIM)*, 2010.

Technical Reports

[T1] **Y. Zhang** and MI. Jordan. Splash: User-friendly Programming Interface for Parallelizing Stochastic Algorithms. *arXiv:1506.07552*, 2015.

Selected Awards & Honors

2017 Best Paper Award, Conference on Learning Theory (COLT).
2016 Outstanding Reviewer Award, International Conference on Machine Learning (ICML).
2015 Baidu Fellowship (awards 8 PhD students every year worldwide).
2013 Microsoft Research PhD Fellowship Finalist.
2006 Silver Medal in Asian Physics Olympiad.
2006 Gold Medal in Chinese Physics Olympiad (5^{th} among 400,000 participants).

Teaching

Graduate Student Instructor, Introduction to machine learning, UC Berkeley **2015**
Graduate Student Instructor, Randomized algorithms for matrices and data, UC Berkeley **2013**

Service

Journal Reviewer: Journal of Machine Learning Research, Annals of Statistics, Mathematical Programming, ACM Transactions on the Web.
Conference Reviewer: ICML (2013 -), NIPS (2013 -), AISTAT (2015 -), IJCAI (2015 -).