

Yi Yang

4545 Center Blvd, Long Island City, NY 11109
<http://yiyangnlp.github.io>

Employment	<i>Senior Research Scientist</i> Bloomberg L.P., New York, NY	Jan. 2017 - present
	<i>Research Intern</i> Machine Learning Department Microsoft Research Redmond, WA	May. 2015 - Aug. 2015 Mentor: Scott Wen-tau Yih
	<i>Research Intern</i> Internet Services Research Center (ISRC) Group Microsoft Research Redmond, WA	May. 2014 - Aug. 2014 Mentor: Ming-Wei Chang
	<i>Research Scientist Intern</i> Speech Group & NLU Group Amazon, inc. Seattle, WA	May. 2013 - Jul. 2013 & May. 2012 - Aug. 2012 Mentor: Bjoern Hoffmeister & Imre Kiss
Education	<i>Doctor of Philosophy</i> School of Interactive Computing Georgia Institute Technology, Atlanta, GA	Aug. 2011 - Dec. 2016 Advisor: Jacob Eisenstein
	<i>Bachelor of Science</i> School of Software Beihang University, Beijing, China	Sep. 2007 - Jun. 2011 GPA: 3.70/4.0, Rank: 2/144
Publications	Yi Yang , Ozan Irsoy, and Kazi Shefaet Rahman. Collective Entity Disambiguation with Structured Gradient Tree Boosting. <i>To appear in the North American Chapter of the Association for Computational Linguistics (NAACL), 2018</i> (Full Paper)	
	Yi Yang and Jacob Eisenstein. Overcoming Language Variation in Sentiment Analysis with Social Attention. <i>Transactions of the Association for Computational Linguistics (TACL), 2017</i> (Journal Paper) (Presented in ACL 2017)	
	Yi Yang , Ming-Wei Chang, and Jacob Eisenstein. Toward Socially-Infused Information Extraction: Embedding Authors, Mentions, and Entities. <i>In Proceedings of Empirical Methods in Natural Language Processing (EMNLP), 2016</i> (Full Paper)	
	Yi Yang and Jacob Eisenstein. Part-of-Speech Tagging for Historical English. <i>In Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL), 2016</i> (Full Oral Paper)	
	Yi Yang and Jacob Eisenstein. Putting Things in Context: Community-specific Embedding Projections for Sentiment Analysis. Technical Report, 2015	
	Yi Yang , Wen-tau Yih, and Christopher Meek. WikiQA: A Challenge Dataset for Open-Domain Question Answering. <i>In Proceedings of Empirical Methods in Natural Language Processing (EMNLP), 2015</i> (Short Oral Paper)	
	Yi Yang and Ming-Wei Chang. S-MART: Novel Tree-based Structured Learning Algorithms Applied to Tweet Entity Linking. <i>In Proceedings of the Association for Computational Linguistics (ACL), 2015</i> (Full Oral Paper)	

Yi Yang and Jacob Eisenstein. Unsupervised Multi-Domain Adaptation with Feature Embeddings. *In Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL), 2015* (Full Paper)

Yi Yang and Jacob Eisenstein. Unsupervised Domain Adaptation with Feature Embeddings. *In Proceedings of the International Conference on Learning Representations (ICLR), 2015* (Workshop Paper)

Yi Yang and Jacob Eisenstein. Fast Easy Unsupervised Domain Adaptation with Marginalized Structured Dropout. *In Proceedings of the Association for Computational Linguistics (ACL), 2014*. (Short Paper)

Yi Yang and Jacob Eisenstein. A Log-Linear Model for Unsupervised Text Normalization. *In Proceedings of Empirical Methods in Natural Language Processing (EMNLP), 2013*. (Full Oral Paper)

Yi Yang and Lambert Mathias. A Statistical Machine Translation Approach to Entity Recognition. *In Proceedings of Amazon Conference on Machine Learning, 2013*

Fangtao Li, Minlie Huang, **Yi Yang**, and Xiaoyan Zhu. Learning to Identify Review Spam. *In Proceedings of the International Joint Conference On Artificial Intelligence (IJCAI), 2011*. (Full Oral Paper)

Minlie Huang, **Yi Yang**, and Xiaoyan Zhu. Quality-biased Ranking of Short Texts in Microblogging Services. *In Proceedings of the International Joint Conference on Natural Language Processing (IJCNLP), 2011* (Full Oral Paper)

Rezarta Islamaj Doan, **Yi Yang**, Aurlie Nvol, Minlie Huang, and Zhiyong Lu. Identifying Protein-protein Interactions in Biomedical Text Articles. *In Proceedings of the BioCreative Challenge Evaluation Workshop, 2010*

Selected Talks

Conference Talk, July 2017. “Overcoming Language Variation in Sentiment Analysis with Social Attention.” *ACL 2017*

Invited Talk, June 2016. “Toward Robust Text Analysis for Language Variation” *University of Pennsylvania; Bloomberg L.P.; Baidu Research; Google Research*

Conference Talk, June 2016. “Part-of-Speech Tagging for Historical English.” *NAACL 2016*

Invited Talk, May 2016. “Efficiently Learning and Applying Dense Feature Representations for Natural Language Processing” *IBM research; Allen Institute for Artificial Intelligence (AI2)*

Guest lecture, December 2015. “Domain Adaptation for Natural Language Processing.” *CS 4650/7650, Natural Language Processing*

Conference Talk, July 2015. “S-MART: Novel Tree-based Structured Learning Algorithms Applied to Tweet Entity Linking.” *ACL 2015*

Guest lecture, March 2015. “Deep Learning of Representations for Domain Adaptation.” *CS8803DL, Deep Learning for Perception*

Intern Talk, August 2014. “Improved Tweet Entity Linking with Structured Gradient Tree Boosting.” *NLP/ISRC group, MSR*

Conference Talk, October 2013. “A Log-Linear Model for Unsupervised Text Normalization.” *EMNLP 2013*

Intern Talk, August 2012. “A statistical machine translation approach to entity recog-

dition.” *Amazon Tech Talk series*

Teaching Experience	<ul style="list-style-type: none">• Computational Journalism• Intro to Enterprise Computing• J2EE Architecture• Java Programming Language• Algorithm Analysis & Design	Spring 2016 Spring 2013 Spring 2011 Fall 2010 Fall 2009
Computer Skills	<ul style="list-style-type: none">• Programming Language: Python, C++, Java, Matlab, C#, JavaScript, HTML• Databases: SQL Server, Oracle, MySql• Tools: Theano, Hadoop, Mallet, Indri, Lucene	
Professional Service	<ul style="list-style-type: none">• Area Co-chair: ACL 2018, NAACL 2018• Program Committee Member<ul style="list-style-type: none">– 2018: EMNLP, RepL4NLP, SocialNLP, WNUT– 2017: ACL, EMNLP, EACL, IJCNLP, DeepStruct, RepL4NLP, SocialNLP, WNUT– 2016: ACL, SocialNLP, WNUT, Human-Computer QA– 2015: ACL, EMNLP– 2014: ACL• Journal Review: Journal of Artificial Intelligence Research• Conference Review: AAAI 2018, ICML 2017, CONLL 2014, EMNLP 2013, ICDCS 2012, WWW 2011	
Selected Honors	<ul style="list-style-type: none">• ACL Student Volunteer• Georgia Tech GSC Travel Grant• ACL Outstanding Reviewer• NAACL Student Volunteer• ACL Student Travel Awards• EMNLP Student Volunteer• Excellent Student of Beihang University (Top 1%)• First Prize for Academic Excellent Student (Top 2%), Beihang University• Ranked 7th in ACM Programming Contest, Beihang University• First Prize for Innovative Student (Top 2%), Beihang University• National Scholarship (Top 2%), P.R.China	2015 2015 2015 2015 2014 2013 2010 2010 2009 2009 2009