Yi Yang

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

Employment

Senior Research Scientist — Jan. 2017 - present
Bloomberg L.P., New York, NY

Research Intern — May. 2015 - Aug. 2015
Machine Learning Department — Mentor: Scott Wen-tau Yih
Microsoft Research Redmond, WA

Research Intern — May. 2014 - Aug. 2014
Internet Services Research Center (ISRC) Group — Mentor: Ming-Wei Chang
Microsoft Research Redmond, WA

Research Scientist Intern — May. 2013 - Jul. 2013 & May. 2012 - Aug. 2012
Speech Group & NLU Group — Mentor: Bjoern Hoffmeister & Imre Kiss
Amazon, inc. Seattle, WA

Education

Doctor of Philosophy — Aug. 2011 - Dec. 2016
School of Interactive Computing — Advisor: Jacob Eisenstein
Georgia Institute Technology, Atlanta, GA

Bachelor of Science — Sep.2007 - Jun.2011
School of Software — GPA: 3.70/4.0, Rank: 2/144
Beihang University, Beijing, China

Publications

Yi Yang and Jacob Eisenstein. Overcoming Language Variation in Sentiment Analysis with Social Attention. *Transactions of the Association for Computational Linguistics (TACL), 2017* (Journal Paper)

Yi Yang, Ming-Wei Chang, and Jacob Eisenstein. Toward Socially-Infused Information Extraction: Embedding Authors, Mentions, and Entities. *In Proceedings of Empirical Methods in Natural Language Processing (EMNLP), 2016* (Full Paper)

Yi Yang and Jacob Eisenstein. Part-of-Speech Tagging for Historical English. *In Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL), 2016* (Oral Full 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. *In Proceedings of Empirical Methods in Natural Language Processing (EMNLP), 2015* (Short Oral Paper)

Yi Yang and Ming-Wei Chang. S-MART: Novel Tree-based Structured Learning Algorithms Applied to Tweet Entity Linking. *In Proceedings of the Association for Computational Linguistics (ACL), 2015* (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

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 recognition." *Amazon Tech Talk series*

Teaching Experience

• Computational Journalism	Spring 2016
• Intro to Enterprise Computing	Spring 2013
• J2EE Architecture	Spring 2011
• Java Programming Language	Fall 2010
• Algorithm Analysis & Design	Fall 2009

Computer Skills

- **Programming Language:** Python, C++, Java, Matlab, C#, JavaScript, HTML
- **Databases:** SQL Server, Oracle, MySql
- **Tools:** Theano, Hadoop, Mallet, Indri, Lucene

Professional Service

- Program Committee Member
  - 2017: ACL, EACL, RepL4NLP, SocialNLP, WUNT
  - 2016: ACL, SocialNLP, WNUT, Human-Computer QA
  - 2015: ACL, EMNLP
  - 2014: ACL
- Conference Review: CONLL 2014, EMNLP 2013, ICDCS 2012, WWW 2011

Selected Honors

- ACL Student Volunteer 2015
- Georgia Tech GSC Travel Grant 2015
- ACL Outstanding Reviewer 2015
- NAACL Student Volunteer 2015
- ACL Student Travel Awards 2014
- EMNLP Student Volunteer 2013
- Excellent Student of Beihang University (Top 1%) 2010
- First Prize for Academic Excellent Student (Top 2%), Beihang University 2010
- Ranked 7th in ACM Programming Contest, Beihang University 2009
- First Prize for Innovative Student (Top 2%), Beihang University 2009
- National Scholarship (Top 2%), P.R.China 2009