

RUOCHEN XU

GHC 5511, 4902 Forbes Ave, Pittsburgh, PA ◇ Pittsburgh, Pennsylvania 15213
(412) · 897 · 3386 ◇ ruochenx@cs.cmu.edu

EDUCATION

Carnegie Mellon University

August 2016 - Now

Ph.D. candidate in Language Technologies, School of Computer Science
Research advisor: Professor Yiming Yang
Overall GPA: 3.97/4.3

Carnegie Mellon University

August 2014 - August 2016

M.S. in Language Technologies, School of Computer Science
Research advisor: Professor Yiming Yang
Overall GPA: 3.96/4.3

Tsinghua University

August 2010 - July 2014

B.S. in Electronic Engineering
Outstanding graduate of Tsinghua University
Scholarship for comprehensive merit in Tsinghua University
Scholarship for excellent academic performance in Tsinghua University
Overall GPA: 3.78/4.0 Rank: 18/210

Washington University in St. Louis

August 2012 - December 2012

Exchange Student in School of Engineering
Scholarship for excellent undergraduate exchange student (sponsor: China Scholarship Council)
Overall GPA: 3.93/4.0

WORK EXPERIENCE

Summer Intern

May 2018 - Aug 2018

supervised by Dr. Tao Ge and Dr. Furu Wei

Microsoft Research Asia

- Built style transfer model for natural language generation

RESEARCH PROJECTS

Analysis of Rare Incident-Event Languages

August 2015 - Now

supervised by Professor Yiming Yang

Carnegie Mellon University

- Ranked at first place at NIST's Low Resource human language technologies (LoReHLT) evaluation
- Addressed resource-scarcity in topic identification task by translating state-of-the-art classification models from high-resource languages to low-resource languages.
- Built neural network text classifier for cross-lingual topic identification with bilingual word embedding and parallel text.
- Utilized word embedding to extend partially observed bilingual dictionary to facilitate model translation.

Machine-Assisted Construction of Knowledge Bases

December 2014 - August 2015

supervised by Professor Yiming Yang

Carnegie Mellon University

- Applied new supervised machine learning algorithms to recommend relevant Web pages to streamline the curation process of knowledge bases.

TAC KBP 2014 Entity Linking Track

supervised by Professor Ji Wu

September 2013 - June 2014

Tsinghua University

- Built end-to-end Entity Linking system with query expansion, knowledge base retrieval, feature extraction and supervised ranking.

Research on Entity Search using Markov Graphical Model

supervised by Professor Kevin C. Chang, July 2013 - September 2013

UIUC

- Designed Tree-structured Markov graphical model to learn the interaction of different feature functions for entity search.

PUBLICATION

Ruochen Xu, Tao Ge, Furu Wei "Formality Style Transfer with Hybrid Textual Annotations" arXiv preprint:1902.08899 (2019)

Ruochen Xu, Yiming Yang, Naoki Otani, Yuexin Wu "Unsupervised Cross-lingual Transfer of Word Embedding Spaces" Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. (EMNLP 2018)

Aldrian Obaja Muis, Naoki Otani, Nidhi Vyas, **Ruochen Xu**, Yiming Yang, Teruko Mitamura, Eduard Hovy "Low-resource Cross-lingual Event Type Detection in Documents via Distant Supervision with Minimal Efforts" Proceedings of the 27th International Conference on Computational Linguistics. (COLING 2018)

Littell, P., Tian, T., **Xu, R.**, Sheikh, Z., Mortensen, D., Levin, L., Tyers, F., Hayashi, H., Horwood, G., Sloto, S., Tagtow, E., Black, A., Yang, Y., Mitamura, T., and Hovy, E. "The ARIEL-CMU situation frame detection pipeline for LoReHLT16: a model translation approach" Machine Translation 32.1-2 (2018): 105-126

Ruochen Xu, Yiming Yang "Cross-lingual Distillation for Text Classification" 55th annual meeting of the Association for Computational Linguistics (ACL 2017)

Ian Gorton, **Ruochen Xu**, Yiming Yang, Hanxiao Liu, Guoqing Zheng "Experiments in Curation: Towards Machine-Assisted Construction of Software Architecture Knowledge Bases" IEEE International Conference on Software Architecture (ICSA 2017)

Andrew Hsi, Yiming Yang, Jaime G. Carbonell, **Ruochen Xu**: "Leveraging Multilingual Training for Limited Resource Event Extraction." Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics. (COLING 2016)

Ruochen Xu, Yiming Yang, Hanxiao Liu, and Andrew Hsi. "Cross-lingual Text Classification via Model Translation with Limited Dictionaries." In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, pp. 95-104. ACM, (CIKM 2016)

PEER REVIEWS

Reviewer:

NAACL 2019; COLING 2018; SIGIR 2017;

IEEE Access; ACM Transactions on Information Systems;

LANGUAGES

Computer Languages

Python, C/C++, Java, JavaScript, MATLAB

Natural Languages

Chinese (native)

English (proficient)