

Xiao Liu

No. 5 South Zhongguancun Street, Haidian District, Beijing, China, 100081
+86 18610789560 • xiaoliu@bit.edu.cn • xiaoliubit.github.io • Google Scholar • GitHub

RESEARCH INTERESTS

Information Extraction; Text Mining; Event Extraction & Schema Induction; Knowledge-boostered Application; Application of NLP & Machine Learning

EDUCATION

Beijing Institute of Technology, Beijing, China

- **Ph.D. Candidate**, Computer Science Sep 2017 – Present
 - Adviser: *Prof. Heyan Huang*
 - Topic: *Event extraction on low-resources domains, including automatically inducing event schemas using unsupervised approaches and jointly extracting multiple events in texts*
- **M.S.**, Computer Science (Rank 5/192) Sep 2016 – Jun 2017
 - Adviser: *Prof. Heyan Huang*
 - *Successive postgraduate and doctoral program*
- **B.S.**, Computer Science (Rank 4/220) Sep 2012 – Jun 2016
 - *Outstanding Graduates at BIT and in Beijing*

Singapore University of Technology and Design, Singapore

- **Joint Ph.D. student**, Information Science Technology and Design Aug 2018 – Mar 2019
 - Host: *Prof. Yue Zhang* and *Prof. Jun Sun*
 - Topic: *Information extraction using neural latent variable models and variational inferences*

Nanyang Technological University, Singapore

- **Joint Ph.D. student**, School of Computer Science and Engineer Mar 2019 – Sep 2019
 - Host: *Prof. Erik Cambria* and *Prof. Yue Zhang*
 - Topic: *Event applications, knowledge-based sentiment analysis and stock market prediction*

EXPERIENCE

Microsoft Research Asia, Beijing, China

- **Research Intern**, Data Mining and Enterprise Intelligence Group Dec 2016 – Nov 2017
 - Worked with Dr Jun Yan and Yaobo Liang on semantic similarity computation and generation.
 - Participated in the FAQ Retrieval project for Pfizer that provides every input question with an answer by retrieving similar standard questions with well-written answers.
 - Built a high-recall filter for top-n similar question-question pairs.
 - Built pair-wise CNNs to classify the question intents for boosting question matching.
 - Built paraphrase generation models by synonyms replacement with language models and beam search.
 - Built a deep learning paraphrase generation model based on an NMT seq2seq framework.

Netease Youdao, Beijing, China

- **Research and Development Intern**, Machine Translation Group Nov 2015 – Dec 2016
 - Improved the result of the Japanese-to-Chinese machine translation service by 4% point increment in BLEU.
 - Cleaned corpus from different sources, optimized word segmentation and ordering strategies, enlarged bilingual and monolingual corpus, trained larger language models, and deployed them into the online production environment.
 - Designed the framework and built the first version of parallel sentence pairs crawler with switchable alignment methods on Spark clusters.

PROFESSIONAL SERVICES	Program Committee Member / Reviewer	
	<ul style="list-style-type: none"> ▪ Conferences: <i>ACL 2020, EMNLP 2020</i> 	
	Secondary Reviewer	
	<ul style="list-style-type: none"> ▪ Conferences: <i>ACL 2019, EMNLP 2018, ACL 2018, COLING 2018, CCL 2017, NLPCC 2016</i> 	
PUBLICATION	<p>[1] Xiao Liu, Heyan Huang, Yue Zhang. Open Domain Event Extraction Using Neural Latent Variable Models. <i>In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)</i>, Florence, Italy, July 28 - August 2, 2019. Long Oral</p> <p>[2] Xiao Liu, Zhunchen Luo, Heyan Huang. Jointly Multiple Events Extraction via Attention-based Graph Information Aggregation. <i>In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)</i>, Brussels, Belgium, October 31 - November 4, 2018. Long Oral</p> <p>[3] Qingkai Min, Libo Qin, Zhiyang Teng, Xiao Liu, Yue Zhang. Dialog State Tracking Using Neural Latent Variable Models. <i>In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI)</i>, 2020.</p> <p>[4] Changsen Yuan, Heyan Huang, Chong Feng, Xiao Liu, Xiaochi Wei. Distant Supervision for Relation Extraction with Linear Attenuation Simulation and Non-IID Relevance Embedding. <i>Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence (AAAI)</i>, Honolulu, Hawaii, USA, January 27-February 1, 2019.</p> <p>[5] Zhunchen Luo, Xiao Liu. Real-time Scholarly Retweeting Prediction System. <i>In Proceedings of the 27th International Conference on Computational Linguistics (COLING)</i>, Santa Fe, New Mexico, USA, August 20-26, 2018.</p>	
SELECTED OPEN-SOURCE PROJECTS	★ represents the number of github stars.	
	<ul style="list-style-type: none"> ▪ JMEE ★138 https://github.com/lx865712528/EMNLP2018-JMEE A neural sequence labeling framework with graph convolution network layer to jointly extract multiple events in a sentence. It gives the best performance on the event extraction task of the ACE 2005 dataset in 2018. ▪ ODEE ★57 https://github.com/lx865712528/ACL2019-ODEE A neural-extended unsuoversised schema induction and open domain event extraction framework which supports incrementally learning. A business report corpus is published along with this framework. 	
TEACHING	Beijing Institute of Technology , Beijing, China	
	<ul style="list-style-type: none"> ▪ Teaching Assistant Sep 2014 – Jan 2015 <ul style="list-style-type: none"> • Algorithm and Data Structure ▪ Teaching Assistant Sep 2017 – Nov 2017 <ul style="list-style-type: none"> • Selected Topics In New Technology for Computer 	
HONORS AND AWARDS	High-level Joint Doctoral Students Scholarship , China Aug 2018 – Sep 2019	
	<ul style="list-style-type: none"> ▪ China Scholarship Council No. 201806030142 ▪ Fully funded high-level Ph.D. students for jointly studying outside China offically. 	
	Beijing Outstanding Graduates , Beijing Jun 2016	
	<ul style="list-style-type: none"> ▪ This award program recognizes undergraduate students in Beijing universities who show outstanding grades and potential in their studies. 	
	Outstanding Students Award , BIT 2012 – 2019	
	<ul style="list-style-type: none"> ▪ Six times. 	

	<ul style="list-style-type: none"> ▪ Beijing Institute of Technology Annual Award and Honor. For rewarding excellent (top 5%) students in BIT in the academic year. 	
	China Century Group Scholarship	2014
	<ul style="list-style-type: none"> ▪ Sponsor and award 4 undergraduate students and 2 postgraduate students who major in computer science every year in BIT by China Century Group. 	
	First-class Ph.D. Academic Scholarship, BIT	2017 – 2019
	<ul style="list-style-type: none"> ▪ Two times. ▪ This award program recognizes top 5 Ph.D. students majored in computer science every year. 	
	First-class M.S. Academic Scholarship, BIT	2016
	<ul style="list-style-type: none"> ▪ This award program recognizes top 30 master students majored in computer science every year. 	
	First-class Undergraduate Academic Scholarship, BIT	2013 – 2016
	<ul style="list-style-type: none"> ▪ Four times. ▪ As for undergraduate student, select top 5% by last semester's GPA. 	
COMPETITIONS AND HACKATHONS	Gold Medal in CCPC Regional	Oct 2016
	<ul style="list-style-type: none"> ▪ The 2nd China Collegiate Programming Contest Regional Hefei Onsite ▪ The rules and difficulties are the same to the ACM/ICPC Regional contests. 	
	Bronze Medal in ACM/ICPC Regional	Oct 2014
	<ul style="list-style-type: none"> ▪ The 39th ACM/ICPC Regional Beijing Onsite 	
	Bronze Medal in ACM/ICPC Regional	Dec 2014
	<ul style="list-style-type: none"> ▪ The 39th ACM/ICPC Regional Shanghai Onsite 	
ACTIVITIES AND ORGANIZATIONS	Association of Programming and Algorithms, BIT	
	<ul style="list-style-type: none"> ▪ Co-founder 	Sep 2013 – Jun 2014
	<ul style="list-style-type: none"> • Organize and invite programming contest experts to deliver lectures. • Host Freshman Enrollment Programming Contest in the end of the first semester every year. 	
	Annual Programming Contest, BIT	
	<ul style="list-style-type: none"> ▪ Problem Provider and Onsite Contest Judge 	May 2015 – May 2017
	<ul style="list-style-type: none"> • Provided several problems for the contests and verify all the selected problems. • Judged program solutions and did the real time Q&A about the problem descriptions during the contests. 	
PROGRAMMING LANGUAGES	<ul style="list-style-type: none"> ▪ Python, C/C++, L^AT_EX, Java, Scala ▪ PyTorch, Tensorflow, Keras ▪ ... 	