Xinya Du

Contact

Assistant Professor
Department of Computer Science
The University of Texas at Dallas
800 West Campbell Road, Richardson, TX 75080

Email: xinya.du@utdallas.edu
Website: https://xinyadu.github.io
Google Scholar

Research Interests

Natural Language Processing, Computational Linguistics, Machine Learning, Deep Learning

Education

Cornell University, Ithaca, NY, USA — Aug 2016 – Aug 2021
Ph.D. in Computer Science (*M.S.* degree granted in Aug 2019)
Dissertation: Towards More Intelligent Extraction of Information from Documents
Advisor: Claire Cardie

Shanghai Jiao Tong University, Shanghai, China — Sep 2012 – Aug 2016
B.E. in Computer Science and Engineering

Experience

University of Texas at Dallas, Richardson, TX — Aug 2022 – Present
Assistant Professor in Computer Science

University of Illinois at Urbana-Champaign, Champaign, IL — Sep 2021 – Aug 2022
Postdoctoral Research Associate, with Prof. Heng Ji

Google AI, Mountain View, CA — May 2020 – Aug 2020
Research Intern

Allen Institute for Artificial Intelligence, Seattle, WA — Sep 2018 – Dec 2018
Research Intern

Microsoft Research, Redmond, WA — May 2018 – Aug 2018
Research Intern

Publications

[1] **Dynamic Global Memory for Document-level Argument Extraction**
Xinya Du, Sha Li, and Heng Ji
In Annual Meeting of the Association for Computational Linguistics (ACL), 2022.

[2] **Automatic Error Analysis for Document-level Information Extraction**
Aliva Das, **Xinya Du**, Barry Wang, Kejian Shi, Jiayuan Gu, Thomas Porter, Claire Cardie
In Annual Meeting of the Association for Computational Linguistics (ACL), 2022.

[3] **RESIN-11: Schema-guided Event Prediction for 11 Newsworthy Scenarios**
Xinya Du, Zixuan Zhang, Sha Li, Heng Ji and the RESIN team
In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL): System Demonstrations, 2022.

[4] **Template Filling with Generative Transformers**
Xinya Du, Alexander M. Rush, and Claire Cardie
In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2021.

[5] **GRIT: Generative Role-filler Transformers for Document-level Event Entity Extraction**
Xinya Du, Alexander M. Rush, and Claire Cardie
In Conference of the European Chapter of the Association for Computational Linguistics

(EACL), 2021.

[6] **Few-shot Intent Classification and Slot Filling with Retrieved Examples**
Dian Yu, Luheng He, Yuan Zhang, **Xinya Du**, Panupong Pasupat and Qi Li
In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2021.

[7] **QA-Driven Zero-shot Slot Filling with Weak Supervision Pretraining**
Xinya Du, Luheng He, Qi Li, Dian Yu, Panupong Pasupat and Yuan Zhang
In Annual Meeting of the Association for Computational Linguistics (ACL), 2021.

[8] **Event Extraction by Answering (Almost) Natural Questions**
Xinya Du and Claire Cardie
In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020.

[9] **Improving Event Duration Prediction via Time-aware Pre-training**
Zonglin Yang, **Xinya Du**, Alexander M. Rush and Claire Cardie
In Findings of the Association for Computational Linguistics: (EMNLP), 2020.

[10] **Document-Level Event Role Filler Extraction using Multi-Granularity Contextualized Encoding**
Xinya Du and Claire Cardie
In Annual Meeting of the Association for Computational Linguistics (ACL), 2020.

[11] **Leveraging Structured Metadata for Improving Question Answering on the Web**
Xinya Du, Adam Fourney, Robert Sim, Claire Cardie, Paul Bennett and Ahmed Hassan Awadallah
In Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL/IJCNLP), 2020.

[12] **Be Consistent! Improving Procedural Text Comprehension using Label Consistency**
Xinya Du, Bhavana Dalvi, Niket Tandon, Antoine Bosselut, Wen-tau Yih, Peter Clark, Claire Cardie
In Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), 2019.

[13] **Harvesting Paragraph-Level Question-Answer Pairs from Wikipedia**
Xinya Du and Claire Cardie
In Annual Meeting of the Association for Computational Linguistics (ACL), 2018.

[14] **Identifying Where to Focus in Reading Comprehension for Neural Question Generation**
Xinya Du and Claire Cardie
In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2017.

[15] **Learning to Ask: Neural Question Generation for Reading Comprehension**
Xinya Du, Junru Shao and Claire Cardie
In Annual Meeting of the Association for Computational Linguistics (ACL), 2017.
Featured in *New Scientist ("Inquisitive bot asks questions to test your understanding") [Link], and*
TechRepublic ("How researchers trained one AI system to start asking its own questions") [Link].

[16] **Cornell Belief and Sentiment System at TAC 2016**
Vlad Niculae, Kai Sun, Xilun Chen, Yao Cheng, **Xinya Du**, Esin Durmus, Arzoo Katiyar, Claire Cardie
In Text Analysis Conference (TAC), 2016.

Awards& Honors

Spotlight Rising Star in Data Science (NLP track) — University of Chicago, 2020
Outstanding Graduate of SJTU — Shanghai Jiao Tong University, 2016
National Scholarship — Shanghai Jiao Tong University, 2013

Teaching Experience

Natural Language Processing, Cornell University, Fall 2019
Teaching Assistant for Prof. Claire Cardie.

Natural Language Processing, Cornell University, Spring 2019
Teaching Assistant for Prof. Yoav Artzi.

Software Engineering, Cornell University, Spring 17, Spring 18
Teaching Assistant for Prof. William Arms.

Introduction to Computing Using Python, Cornell University, Fall 2016
Teaching Assistant for Prof. Walker White.

Mentoring Experience

Zonglin Yang, Cornell CS MEng student, 2020 – 2022.
Topic: Commonsense and case-based reasoning for NLP.
Publications: EMNLP (Findings), 2020.

Barry Wang, Cornell CS undergraduate student, 2021 – 2022
Topic: Automatic error analysis for information extraction.
Publications: ACL 2022, SciNLP 2021.

Aliva Das, Cornell CS undergraduate student, 2021 – 2022.
Topic: Automatic error analysis for information extraction.
Publications: ACL 2022, SciNLP 2021.

Maitreyi Chatterjee, Cornell CS undergraduate student, Spring 2021.
Topic: Applying neural document-level IE model to scientific domain.

Rishi Malhotra, Cornell CS undergraduate student, Spring 2021.
Topic: Applying neural document-level IE model to scientific domain.

Professional Services

Journal Reviewer:

IEEE Transactions on Knowledge and Data Engineering (TKDE)

IEEE Transactions on Audio, Speech and Language Processing (TASLP)

ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)

ACM Transactions on Knowledge Discovery from Data (TKDD)

Knowledge and Information Systems (KAIS)

AI Communication

Conference Committee Member:

Annual Meeting of the Association for Computational Linguistics (ACL)

Conference on Empirical Methods in Natural Language Processing (EMNLP)

Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)

Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL/IJCNLP)

International Joint Conference on Artificial Intelligence (IJCAI)

AAAI Conference on Artificial Intelligence (AAAI)

Natural Language Processing and Chinese Computing (NLPCC)

Conference on Computational Natural Language Learning (CoNLL)

Workshop on Noisy User-generated Text (W-NUT)

Workshop on Machine Reading for Question Answering (MRQA)

Joint Conference on Lexical and Computational Semantics

Volunteering Activities:

Member of Cornell CS Department PhD Admission Committee (Year 2021).

Volunteer for Cornell CS Department PhD Visit Day (Year 2019, 2020, 2021).

Student Volunteer for ACL 2017, ACL 2018, EMNLP 2017.

Talks

Towards More Informed Extraction of Events from Documents
In Rising Stars in Data Science Workshop, University of Chicago, Jan 2021.
In Tencent AI Research, Nov 2020.

Event Extraction by Answering (Almost) Natural Question
In UIUC Class Information Extraction and Knowledge Acquisition, UIUC, Sep 2020

LwLL: Progress on the NLP Front
In DARPA site visit (Online), Cornell University, Apr 2020.

Harvesting Paragraph-Level Question-Answer Pairs from Wikipedia
In the 56th Annual Meeting of the Association for Computational Linguistics, July 2018.